A COMBINED STATISTICAL AND DYNAMICAL APPROACH TO REGIONAL FORECAST MODELING OF OPEN OCEAN CURRENTS

A thesis presented

by

Kueishiong Tu

to

The Division of Applied Sciences

in partial fulfullment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Applied Mathematics

Harvard University

Cambridge, Massachusetts

February 1981

PREFACE

I would like to take this opportunity to acknowledge the help I have received from many sources during the preparation of this thesis. I am indebted to my thesis advisor, Professor Robinson for his suggestion of the problems and his scientific guidance throughout, especially in the interpretation of results. I am grateful to my thesis committee, Professors Carrier and Krook and particularly Professor Anderson for his detailed comments; and to Drs. Haidvogel and Miller for their advice and also for taking time to read the thesis. Mr. E. Carter contributed to an early version of the statistical numerical model and much programming support was obtained from Mr. J. Groisser.

The work reported here was supported by the Office of Naval Research (Contract #N00014-76-C-0225 to Harvard University) and the National Aeronautics and Space Administration (Grant #NASA-NSG-5228 to Harvard University). Computations were carried out at the Goddard Space Flight Center Modelling and Simulation Facility (GSFC/GMSF) of the Goddard Laboratory of Atmospheric Sciences (GLAS).

Last, but not least, I have to thank my parents for bringing me up and taking pains in educating me through. Thanks are also due to Harvard University for the early supports of my graduate study. I thank John Deep for proofreading the first draft and Marla Pinkston for her excellent typing of the thesis.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

SYNOPSIS

This thesis presents various approaches to investigate the local dynamical processes and to forecast the evolution of the mesoscale currents in an open block of ocean -- the dynamical approach, the statistical approach, and the combined statistical and dynamical approach. The dynamical approach uses a finite element numerical model to solve the linearly damped barotropic vorticity equation. An exterior dynamical calculation in a 1000 x 1000 square km domain is performed to generate a simulated open ocean data set. This is to provide initial, boundary and verification data for future forecast simulation experiments over an interior 500 x 500 square km domain. The well-posedness of the open boundary conditions are treated next. Here we extend Sundström work (23) to consider boundary disturbances also. The same treatment for the quasi-geostrophic potential vorticity equation is first presented here. Some special cases of estimating the errors of the solutions in terms of the errors of the initial and boundary data in the studies of boundary conditions are given in which the error bounds are exact. Some benchmark dynamical calculations are performed for later comparison purposes.

The statistical approach uses Gandin's objective analysis formula to do the space-time interpolation. Some novelties in the applications of this formula are the correlation function used, which is computed from the simulated ocean data set, and the data points selecting algorithm included, which

is to increase the efficiency and stability of the computation. Several statistical forecast examples using data at the same time level and several time levels are given.

The combined statistical and dynamical approach uses the statistical model to provide the initial and boundary conditions and to assimilate recently available data into the dynamical model. Several examples of using this combined statistical and dynamical approach are given. The question of exploiting one data realization is discussed in depth.

CHAPTER I

INTRODUCTION TO OPEN OCEAN REGIONAL FORECAST MODELLING

I.1 The Problem: Time and Space Scales, Data Types and Potential Availability.

The oceans cover 70 percent of our globe. They, together with the atmosphere, determine the weather and climate of the earth, are abundant in food and mineral resources, are an important avenue of transportation, are a place for military operation, and are now a site for chemical and nuclear waste disposal. Because of their practical importance, there has been a growing interest in the studies of the oceans, which has evolved into four disciplines: physical oceanography, biological oceanography, chemical oceanography and geological oceanography.

The goal of the physical oceanographers is to obtain a systematic quantitative description of the physical characteristics of the ocean waters and of their movements. The former includes such aspects as the temperature and salt content, which determine density and infer vertical movement, and also includes dissolved substances or biological species in so far as they yield information about the currents. The latter include the major ocean currents which circulate continuously, the low frequency, mesoscale eddies in the mid-ocean, the variable coastal currents, the reversing tidal currents, the rise and fall of the tide, and the waves generated by wind or earthquake. These oceanographers also want to understand these motions in terms of basic physical principles--that is, accord-

ing to the general laws of the dynamics and thermodynamics of fluids--and to model them. "Why is there a Gulf Stream? What causes waves to break? When, where and how are energy, heat, and momentum put into the sea? How are these quantities moved about from place to place and depth to depth and ultimately dissipated." (Robinson & Simmons (27))

"The very large size of the oceans makes it difficult to achieve this goal. Man is small, his resources limited, and his technology relatively primitive compared to what is needed. Moreover, large-scale fluid flows are almost always turbulent and the ocean is of no exception. Consequently, there are a wealth of physical phenomena occurring in the ocean, characterized by a variety of time and space scales. The time scale ranges from seconds, hours, days, years to longer than millennia. The space scale ranges from millimeters (tiny surface capillary waves), meters (small-scale turbulence), kilometers (powerful and variable current systems) to the circumference of the earth (global-scale general circulation). Some of the scales are directly imposed by ice ages, or seasonal or daily heating and cooling of the sea surface, or the size of an ocean basin, or the extent of submarine mountain ranges. Other scales arise spontaneously from internal resonances or turbulence."* To obtain useful and believable results in any single attempt towards the final goal, we have to limit the scales of study.

The objects of Study of the Harvard Open Ocean Modelling Group are the dynamics of the low frequency variability of ocean currents (mid-ocean eddies and intense current sys-

* Robinson & Simmons (27)

tems) and the role of these currents in the dynamics and transports of the general circulation which are relevant to regions of the western North Atlantic Ocean. "The western North Atlantic main subtropical gyre is perhaps the most intensively studied relatively large region of the world's oceans. It is known to contain variable current features (meanders, rings, eddies, fronts, lenses, waves) with distinctive synoptic characteristics and a variety of underlying dynamics. Although detailed processes are not yet known, this area is known to have subregions active in eddy energy production, conversion and transmission, substantial heat transport, and vigorous air-sea interactions. It is the scene of a large amount of present and potential practical human activities." (Robinson, private comm.)

The approaches we adopted are: (i) to construct a dynamical and a statistical numerical model for an arbitrary block of open ocean in which the medium-scale turbulence is resolved while the subgrid-scale turbulence is treated implictly and empirically, (ii) validate and evaluate these numerical models by comparing the model results with real ocean data sets. "The real ocean data may be from the conventional data sources such as the temperature, velocity, salinity data collected from hydrographic stations, real time XBT's, SOFAR floats and drifter observations, and satellite infrared sea surface temperature, etc. Potential available data sources include satellite altimetry, acoustic tomography, and satellite monitored mother buoys reporting measurements from moored instrumentation (such

as inverted echo sounders) and continuously recording current and temperature." (Robinson, priv. comm.) (iii) Combine the dynamical and statistical models for the purpose of assimilating the most recently acquired data into the dynamical model in some statistically optimal sense.

The ultimate goal of our studies is to construct a model of the ocean that is a good enough physical analogue of the real ocean so that we are able to use it to investigate the local dynamical processes and to forecast the evolution of the mesoscale currents in a block of ocean.

Because ocean currents transport nutrients, chemicals and particulate matter, as well as heat and energy, knowledge of them is also important to those concerned with the life and resources of the ocean and with waste disposal in the ocean, as well as to those who seek to understand our climate. "Physical oceanography is by no means an isolated scientific discipline; it is interactive and interconnective with meteorology, biological oceanography, chemical oceanography, geological oceanography and environmental sciences." (Robinson & Simmons (27)) Better knowledge of the ocean circulation gained from good ocean circulation numerical models can help us understand more about the climate, the global distribution of the chemicals, and the biological productivity of the sea by coupling the ocean circulation numerical models to the atmospheric, chemical, and biological numerical models. "This effort can benefit the society by contributing to the rational management of the planetary environment and resources." (Robinson (24))

The following sections describe briefly the approaches we have adopted. Detailed descriptions will be presented in the later chapters. The results presented in this thesis are restricted to those obtained from the simulation data. However, they form the starting point and pave the way for working with the real data.

I.2 The Dynamical Model

"The Ocean is a classical fluid system. The basic physical laws governing its dynamic behavior are those of classical hydrodynamics and thermodynamics. They include the conservation of mass, momentum, and energy. These conservation laws can be expressed in equations that are applicable to a continuum. These, together with the equation of state of seawater and the conservation equation of the combined specific density of all the dissolved salts that influence the mass density of the water, constitute a system of partial differential equations which are complete and adequate to describe the evolution of the state of the system." (Robinson (25)) Thus, one way to forecast the evolution of the ocean currents is to solve the initial-boundary value problem posed by the system of partial differential equations. This is the dynamical approach. The predictability in the real cases is much lower than that in the simulated cases due to additional sources of errors such as physical and observational errors besides computational ones.

However, the basic model equations are very general. They contain a wealth of distinct phenomena and are applicable to many special circumstances of fluid flows. For example, "they

contain solutions corresponding to acoustic waves, surface and internal gravity waves and bores. They can describe the breaking of waves on the beaches, the wakes of ships and fish, convective overturning in deep heated trenches and the massive coursing and transient meandering of the Gulf Stream." (Robinson (25)) In order to limit the study to mid-ocean mesoscale eddies, we perform a scale analysis of the original equations. Keeping only those terms which are significantly influenced by the effects of the earth's rotation and ignoring the others, we obtain the quasi-geostrophic potential vorticity equation. It is a good model for thermocline and deep mesoscale eddies in the mid-oceans.

Due to its analytic inaccessibility, numerical methods are used to solve the initial-boundary value problem of the non-linear quasigeostrophic potential vorticity equation. However, because of the speed and storage constraints of the contemporary computers, it is neither feasible nor desirable to model the world ocean. The domain we chose for modelling study is thus limited to a block region in the mid-ocean in order to have the required resolution and accuracy. The boundary of the domain chosen, which is not a solid boundary, is called an artificial boundary or open boundary. In this case, what is an appropriate boundary condition on this open boundary is not so obvious. It is called the "open boundary condition problem." This open boundary condition should provide enough information on the interactions between the eddies in the interior and the exterior flow, and in such a way that the solution inside the domain can be uniquely well determined. To answer the question "what is a

well-posed (existence, uniqueness, continuous dependence) open boundary condition?" turns out to be not so easy. Not aware of all aspects or not taking a broader view of the problem could give rise to misleading conclusions. Take the barotropic vorticity equation, which is the two-dimensional analog of the quasi-geostrophic potential vorticity equation, as an example. The traditional boundary condition, the so-called Charney, Fjörtoft, and von Neumann boundary condition, is to specify the streamfunction at all points at an initial time t_0, the streamfunction at all boundary points and vorticity at only inflow boundary points $\forall\, t > t_0$. In their studies of this boundary condition, Sundström (23) and Bennett (5), who considered only one aspect of the problem, reached contradictory conclusions. Neither of them was completely right or wrong. Sundström, who ignored the existence question, proved that the solution depends continuously on the initial condition. He did not establish the same for the boundary condition. Bennett, who claimed the CFVN boundary condition is ill-posed because a set of complex constraints have to be satisfied by the boundary data at the points where the flow is tangential for the equation to have a smooth solution, is too narrow in his viewpoint. The constraints should be allowed in the formulation of initial-boundary value problems, no matter how complicated they are. Otherwise, almost all nonlinear initial-boundary value problems will be ill-posed, which is undesirable. The status of the "open boundary condition problem" is that we should have no difficulty in determining what boundary information--which represents the interaction between the interior and exterior flow--is

enough for us to determine uniquely the solution inside the domain. But what the minimal, non-redundant boundary information is, or in other words, what constraints this boundary information has to satisfy, for the problem to be well-posed is still unknown. However, this should not constitute a big problem, because sometimes the solution does not have to satisfy the boundary conditions exactly (in some weaker form will suffice) and sometimes, in discrete formulations, the computational boundary needs not even to be the original boundary. Thus, from a practical point of view, we can still carry out the numerical calculations even though the well-posedness of the problem is not well resolved.

The most common numerical methods used to solve the initial-boundary problem are the finite difference, finite element and pseudospectral methods. In Haidvogel et al (12) these three methods are used to solve the barotropic vorticity equation in an open ocean. Integrations of moderate length (5-10 periods of the known analytic solution) are performed to determine the accuracy, stability, and efficiency of each method as a function of problem class and the associated physical and computational non-dimensional parameters. It is found that all three methods are, in general, capable of delivering stable and efficient solutions. However, the finite element and pseudospectral methods are, on the average, 4 to 15 times more accurate respectively than the finite difference method. More precisely, they are not compared from the same ground. The finite difference, finite element, pseudospectral method used in the comparisons are of 2nd, 4th and infinite order of accuracy respectively. The comparisons are based on the same

space and time discretizations. The finite element quasigeostrophic open ocean numerical model has since been adopted as a standard dynamical forecast model. It is used to generate a simulated ocean data set with which various controlled forecast experiments are performed and diagnosed. The details of the dynamical forecast model and some simulation results will be presented in Chapter II.

I.3 The Statistical Model

Another way to forecast the ocean currents is by statistical methods. If we have sufficient statistical knowledge about how the future and the present are correlated, knowing the currents today provides some clue as to what the currents may look like tomorrow. There are many approaches in constructing statistical models. One may use the state-space approach (Box (6)) in which variables of interest are formed as a state vector which is related to the state vector at a previous time through a dynamical equation. The parameters of the dynamical equation are identified from the statistics of the system. The approach we are going to adopt here is the space-time objective analysis. Objective analysis (Gandin (10)) has long been used to map coarsely distributed data into the whole field. Here this technique is extended to do interpolation not only in space but also in time. In general, we can use data at different points and different time levels to interpolate to any point at any time. The interpolation which uses the data at a certain time level to obtain the field in the future forms our statistical forecast model.

The success of our statistical model relies on a good correlation function, because it contains the information on

how the fields are related. Thus, special care has been taken to find a correct and usable correlation function. The approach we followed is to compute the correlation function directly from the simulated ocean data set generated by the dynamical model. Using this computed correlation function, our statistical model has been able to deliver good and credible results.

The statistical model can be applied in various ways other than the pure statistical forecast. For example, it can be used as an interpolation scheme both spatially and temporally to provide initial and/or boundary conditions for the dynamical model. In this way the dynamical and statistical models are combined together. The details of the statistical model and some simulation results will be presented in Chapter III.

I.4 The Combined Statistical and Dynamical Approach

Although the dynamical model has been proven capable of delivering remarkably accurate forecasts if it is provided with perfect initial and boundary data, these data are usually very scarce or even missing in typical mid-ocean conditions. Furthermore, in the open domain case the boundary condition changes with time. It has to be predicted before we can do the dynamical forecast. Also, if there are new data available, they can be assimilated into the model to obtain a better estimate of the dynamical field. It is to provide the necessary initial and boundary conditions and to assimilate recently available data that statistical methods are introduced to the dynamical model.

This is our combined statistical and dynamical approach.

In this approach, the statistics needed to complete the study are usually not known. Especially in the open domain case, the statistics of boundary forcing are very difficult to estimate and predict. The approach we adopt here is to compute these statistics directly. This is possible because of the availability of verification data provided by the simulated dynamical calculation. However, they are not computed on a grid point-by-point basis, but on a region-by-region basis. Within each region, the statistics are treated as uniform. The intention is to replace these computed statistics later by estimates which are averages of the computed statistics from several realizations. It is these estimated statistics that are going to be used in the forecasting studies which employ real ocean data.

Some examples of the combined statistical and dynamical approach are given in Chapter IV. They constitute the acme of this thesis. To our knowledge this is the first serious attempt to apply these techniques to dynamical oceanography. It is our belief that this methodology will make a substantial contribution to understand the dynamics of ocean currents and to describe them via a "network approach" involving a multicomponent real-time observational network and dual numerical model.

I.5 Brief Summary of Thesis Topics and Results

In this section, we briefly summarize the topics contained and results presented in this thesis. In Chapter II, we discuss the dynamical model. The quasigeostrophic potential vorticity equation, which is our fundamental model equation, is first

derived in Section II.1 from a scale analysis of the classical fluid dynamical equation. We then describe the finite element method to solve the linearly damped barotropic vorticity equation, which is a slight variation of the quasigeostrophic potential vorticity equation in two dimensions. The simulation environment is described in Section II.2. The major work of our simulation experiments is performed with the linearly damped barotropic vorticity equation.

The question whether a set of open boundary conditions are well-posed is discussed in Section II.3. We find that the functional classes for the boundary variables involved are not arbitrary. They must be restricted in some related ways for the solution to depend continuously on the initial and boundary data. But the necessary and sufficient condition for an open boundary condition to be well-posed is still unknown today. This extends Sundström's work (23) to consider boundary disturbances also. The same treatment of the quasigeostrophic potential vorticity equation is a new result.

We obtain as a bonus an estimate of the error of the solution as a function of the errors of the initial and boundary conditions accumulated in time. However, this is not a very tight bound. In Section II.4, we obtain an exact bound for the linear barotropic vorticity equation with zero initial condition and perfect boundary condition. These results are also new.

Some results of the persistent dynamical calculation, in which the boundary condition is kept frozen from the initial time, are given in Section II.5. The purpose is to provide a benchmark with which future simulation results can be compared.

In Section II.6, we study how the dynamical forecast depends on the persistent interval and the reinitialization interval. We observe that the boundary condition has a dominant influence on the dynamical forecast. Thus, some time must pass before the reinitialization can be effective. This new result is unexpected.

In Chapter III, we discuss the statistical model. The space-time objective analysis formula is derived in Section III.1 The correlation function to be used and the algorithm for selecting data points are discussed in Section III.2. These are new results. In Section III.3, we study how the statistical forecast depends on the model parameters. The purpose is to choose the parameters properly to obtain the best result. The statistical forecast using data at one time level is employed to perform this study. In Section III.4, the results of statistical forecasts using data at several time levels are given, which may then be used to construct initial conditions. In Chapter IV, we discuss the combined statistical and dynamical approach. The optimal combination of estimates formula is derived in Section IV.1. This formula is used to obtain a better estimate from several a priori estimates. This is the first time that the optimal estimation theory is applied in oceanographic studies. In Section IV.2, we discuss how to reduce the amount of statistical information required to perform the optimal combination. Error models are introduced in which the statistics in each region are treated as uniform. We give some examples of the combined statistical and dynamical approach in Section IV.3. The first example studies how the dynamical

forecast depends on the initial and boundary condition samplings. This may be used to find the most efficient data sampling scheme satisfying a given forecast accuracy requirement. The second example performs the dynamical calculation with the boundary condition obtained from the statistical forecast. The result is better than either the persistent dynamical forecast or the persistent statistical forecast. The third example updates the interior with a patch of data. We find the updating does not do much good if the boundary condition is not updated as well. If we have a perfect boundary condition but an imperfect initial condition, then the interior updating will certainly help. This is identical to the classical updating results. The fourth example combines the persistent statistical forecast and persistent dynamical forecast to obtain a better estimate. In Section IV.4, we explore how to obtain the best forecast from a given perfect data field. It is shown that a good estimate of the boundary condition is very important in obtaining a good result. The last chapter draws some conclusions and future research suggestions.

CHAPTER II
THE DYNAMICAL APPROACH

II.1 Quasigeostrophic Dynamic Model

In this section, we describe the dynamical model in detail. We first derive the quasigeostrophic potential vorticity equation from a scale analysis of the classical fluid dynamical equations. Although this equation has been derived in many places (see Charney (7), for example), we rederive it here in an oceanic context. Then we describe the finite element method used to solve the linearly damped barotropic vorticity equation, which is a slight variation of the quasigeostrophic potential vorticity equation in two dimensions, numerically. The basic settings for all the simulation experiments reported in this thesis are described at the end.

The classical fluid dynamical equations governing the motions of the ocean expressed in a coordinate frame rotating with the earth are

$$u_t + u u_x + v u_y + w u_z - f v + e w = -\frac{p_x}{\rho} \qquad (2.1)$$

$$v_t + u v_x + v v_y + w v_z + f u = -\frac{p_y}{\rho} \qquad (2.2)$$

$$w_t + u w_x + v w_y + w w_z - e u + g = -\frac{p_z}{\rho} \qquad (2.3)$$

$$\rho_t + u \rho_x + v \rho_y + w \rho_z + \rho(u_x + v_y + w_z) = 0 \qquad (2.4)$$

$$T_t + u T_x + v T_y + w T_z = 0 \qquad (2.5)$$

where u is the velocity in the x-direction pointing to the east.

v is the velocity in the y-direction pointing to the west.

w is the velocity in the z-direction pointing vertically upwards.

ρ is the density of the ocean water.

T is the potential temperature.

$f = 2\Omega \sin\theta \qquad e = 2\Omega \cos\theta$

Ω is the earth's angular rotating rate.

θ is the latitude.

g is the gravity constant.

Making the hydrostatic approximation, equation (2.3) is replaced by

$$g = -\frac{P_z}{\rho} \tag{2.3a}$$

The Boussinesq approximation is used, which treats the density as constant except when it is associated with the gravity where a buoyancy force is introduced due to density difference.

Equations (2.3a) and (2.4) become

$$\rho g [1 - \alpha (T - T_s(z))] = -P_z \tag{2.3b}$$

$$u_x + v_y + w_z = 0 \tag{2.4a}$$

respectively.

Introduce the following scale amplitudes:

$$(x, y) = L(x', y') \qquad z = Dz' \tag{2.6a}$$

$$(u, v) = U(u', v') \qquad w = U\frac{D}{L} w' \tag{2.6b}$$

$$t = \frac{L}{U} t' \tag{2.6c}$$

$$P = -\rho g z + \rho f_0 U L P' \tag{2.6d}$$

$$f = f_0 (1 + \hat{\beta}\frac{L}{a} y') \tag{2.6e}$$

$$T = T_s(z) + \frac{f_0 L U}{\alpha g D} T' \tag{2.6f}$$

$$S(z) = \frac{dT_s(z)}{dz} = \frac{N^2}{\alpha g} S'(z) \tag{2.6g}$$

where L is the characteristic horizontal scale.

D is the characteristic vertical scale.

U is the characteristic horizontal velocity.

f_0 is the Coriolis parameter at some mid-latitude .

a is the radius of the earth.

$T_s(z)$ is the basic stratified temperature field.

N is the Brünt-Vaisala frequency, which is equal to $\left(\frac{g}{T_s}\frac{dT_s}{dz}\right)^{1/2}$

α is the thermal expansion coefficient.

The dynamical equations in non-dimensional form are

$$R_o(u_t + uu_x + vu_y + wu_z) - (1+\hat{\beta}\frac{L}{a}y)v + \frac{g}{f_0}\frac{D}{L}w = -p_x \qquad (2.7)$$

$$R_o(v_t + uv_x + vv_y + wv_z) + (1+\hat{\beta}\frac{L}{a}y)u = -p_y \qquad (2.8)$$

$$T = p_z \qquad (2.9)$$

$$u_x + v_y + w_z = 0 \qquad (2.10)$$

$$R_o(T_t + uT_x + vT_y + wT_z) + \beta S(z)w = 0 \qquad (2.11)$$

where the prime's have been dropped and

$$R_o = \frac{U}{f_0 L} \qquad \text{(Rossby number)} \qquad (2.12)$$

$$\beta = \frac{N^2 D^2}{f_0^2 L^2} \qquad \text{(Burger number)} \qquad (2.13)$$

For the scale of motions which are significantly influenced by the effects of the earth's rotation, the following assumptions are made:

$$R_o \ll 1 \qquad (2.14)$$

$$\beta = O(1) \qquad (2.15)$$

$$\hat{\beta}\frac{L}{a} = O(R_o) \qquad (2.16)$$

$$\frac{D}{L} \ll 1 \qquad (2.17)$$

Expand the flow variables by the following perturbation approximation

$$A = A_0 + R_0 A_1 + \cdots \tag{2.18}$$

The zeroth-order equations are

$$-v_0 = p_{0x} \tag{2.19}$$

$$u_0 = p_{0y} \tag{2.20}$$

$$T_0 = p_{0z} \tag{2.21}$$

$$u_{0x} + v_{0y} + w_{0z} = 0 \tag{2.22}$$

$$\mathcal{B}\, S(z)\, w_0 = 0 \tag{2.23}$$

Equations (2.19) and (2.20) are the so-called geostrophic relations. From (2.23)

$$w_0 = 0 \tag{2.23a}$$

So (2.22)

$$u_{0x} + v_{0y} = 0 \tag{2.22a}$$

Thus, the zeroth-order motions are horizontally non-divergent and have no vertical components.

The first-order equations are

$$u_{0t} + u_0 u_{0x} + v_0 u_{0y} - v_1 - \beta y v_0 = -p_{1x} \tag{2.24}$$

$$v_{0t} + u_0 v_{0x} + v_0 v_{0y} + u_1 + \beta y u_0 = -p_{1y} \tag{2.25}$$

$$T_1 = p_{1z} \tag{2.26}$$

$$u_{1x} + v_{1y} + w_{1z} = 0 \tag{2.27}$$

$$T_{0t} + u_0 T_{0x} + v_0 T_{0y} + B\, S(z)\, w_1 = 0 \tag{2.28}$$

where

$$\beta = \frac{\hat{\beta}}{R_0}\,\frac{L}{a} \tag{2.29}$$

From (2.24), (2.25), (2.27), (2.28) and (2.21), the following equation is derived

$$\left(\frac{\partial}{\partial t} - \psi_y \frac{\partial}{\partial x} + \psi_x \frac{\partial}{\partial y}\right) q = 0 \tag{2.30}$$

where

$$\psi = p_0 \tag{2.31}$$

$$q = \nabla_H^2 \psi + \frac{1}{\beta}\left(\frac{\psi_z}{\sigma(z)}\right)_z + \beta y \tag{2.32}$$

$$\nabla_H^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \tag{2.33}$$

After we have solved for ψ, we can find the zeroth-order horizontal velocity, which is our main interest, from (2.19) and (2.20)

$$u_0 = -\psi_y \tag{2.20a}$$

$$v_0 = \psi_x \tag{2.19a}$$

This is the quasigeostrophic potential vorticity equation. The potential vorticity q along the projection of a particle path on a horizontal plane is conserved. This equation played a significant role in the early development of numerical weather forecasting.

Although the numerical model used to solve the quasigeostrophic potential vorticity equation has been implemented and successfully tested and validated, we shall limit the following discussions of numerical solution and simulation experiments to the linearly damped barotrophic vorticity equation only, which is a slight variation of the quasigeostrophic potential vorticity equation in two dimensions (with bottom friction included).

The linearly damped barotropic vorticity equation on a β-plane are (in dimensional form):

$$\left\{\frac{\partial}{\partial t} + J(\psi, \cdot)\right\}(\zeta + f) = -K\zeta \tag{2.34}$$

$$\zeta = \nabla^2\psi \tag{2.35}$$

where

$$f = f_0 + \beta y \tag{2.36}$$

$$J(\psi, \cdot) = \frac{\partial \psi}{\partial x}\frac{\partial}{\partial y} - \frac{\partial \psi}{\partial y}\frac{\partial}{\partial x} \tag{2.37}$$

$$(u, v) = (-\psi_y, \psi_x) \tag{2.38}$$

K is the bottom friction coefficient.

Using the scaling

$$\{x, y, t, \psi, \zeta\} = \left\{dx', dy', \frac{t'}{\beta d}, V_0 d\psi', \frac{V_0}{d}\zeta'\right\} \tag{2.39}$$

we have (omitting the prime's)

$$\left\{\frac{\partial}{\partial t} + \varepsilon J(\psi, \cdot)\right\}\zeta + \psi_x = -\kappa\zeta \tag{2.40}$$

$$\nabla^2\psi = \zeta \tag{2.41}$$

where

$$\varepsilon = \frac{V_0}{\beta d^2} \qquad (\beta\text{-Rossby number}) \tag{2.42}$$

$$\kappa = \frac{K}{\beta d} \qquad \text{(bottom frictional Ekman number)} \tag{2.43}$$

ε is a measure of nonlinearity and κ is a measure of dissipativity.

The Harvard Open Ocean Dynamical Model used for dynamical forecast simulation studies is a finite element model. This model uses the finite element method to solve equations (2.40) and (2.41) in an arbitrary rectangular region. For more detailed

information on applying the finite element numerical method to ocean modellings, see Hirsh (13).

A computational grid domain used in the finite element model is shown in Figure 2.0. The bases are chosen to be the bilinear elements which are products of two linear elements in variable x and y respectively, i.e.

$$\phi_{ij}(x,y) = \phi_i(x)\,\phi_j(y) \tag{2.44}$$

where $\phi_i(x)$ is a piecewise linear function which is equal to 1 at $x=x_i$ and 0 at other grid points. Each ϕ_{ij} is a pagoda function.

The time dimension is treated by the following second-order time-differencing scheme.

$$\zeta^{n+1} = \zeta^n - \Delta t\left\{\frac{3}{2}Q^n - \frac{1}{2}Q^{n-1}\right\} - \frac{\alpha\Delta t}{2}(\zeta^{n+1}+\zeta^n)$$

$$\zeta^{n+1} = \frac{(1-\frac{\alpha\Delta t}{2})}{(1+\frac{\alpha\Delta t}{2})}\zeta^n - \frac{\Delta t}{(1+\frac{\alpha\Delta t}{2})}\left\{\frac{3}{2}Q^n - \frac{1}{2}Q^{n-1}\right\} \tag{2.45}$$

where

$$Q^n = \varepsilon J(\psi^n, \zeta^n) + \psi_x^n \tag{2.46}$$

Let

$$\psi^n(x,y) = \sum_{i,j}\psi_{ij}^n\,\phi_i(x)\phi_j(y) \tag{2.47}$$

$$\zeta^n(x,y) = \sum_{i,j}\zeta_{ij}^n\,\phi_i(x)\phi_j(y) \tag{2.48}$$

Substitute (2.48) and (2.49) into (2.46). Multiplying by a basis function $\phi_p(x)\phi_q(y)$ and integrating over the entire domain (this is using the Galerkin's principle), we get

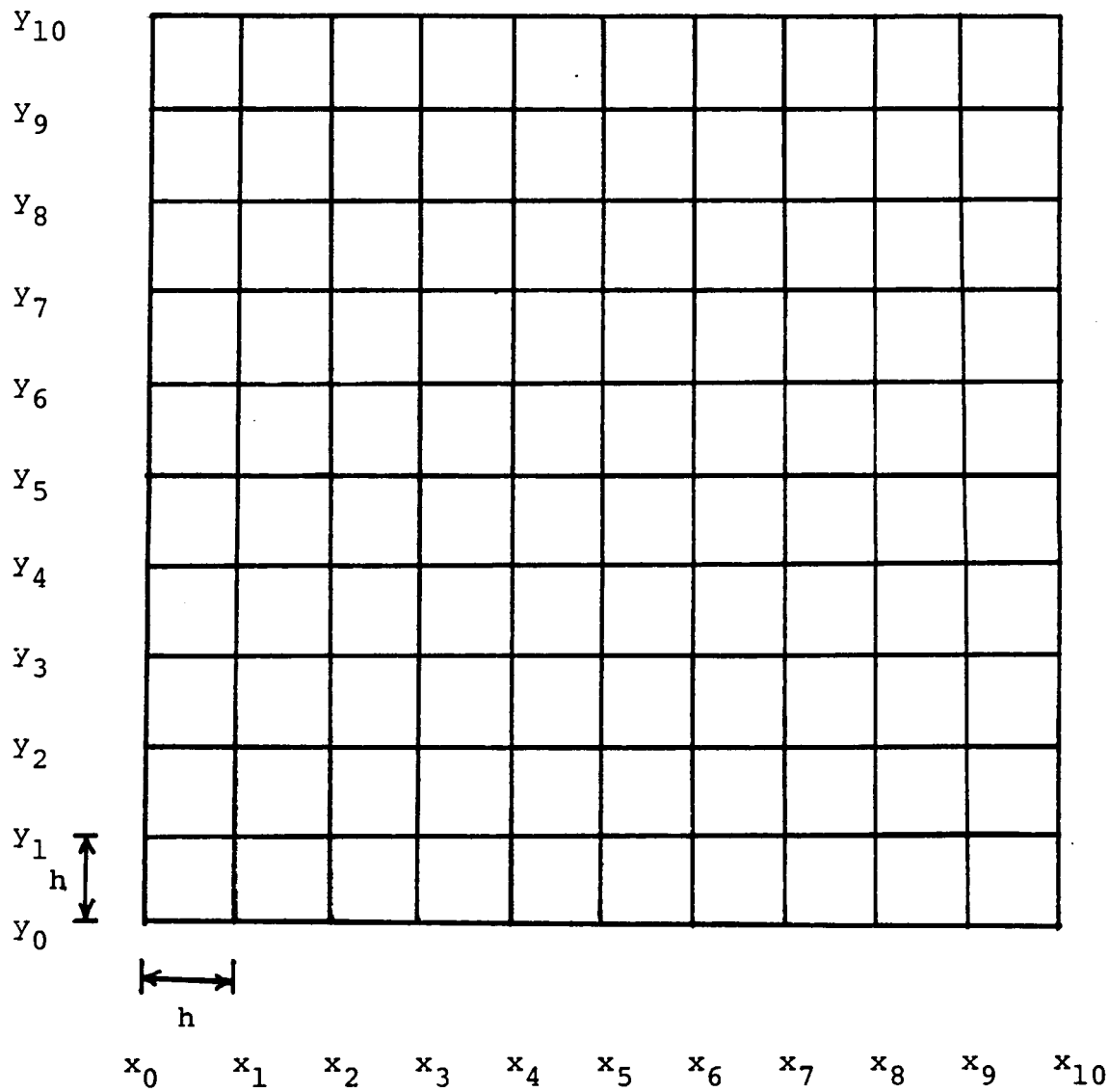


Figure 2.0 A computational grid domain for the finite element dynamical numerical model.

$$M(\zeta^{n+1}) = \frac{(1-\frac{\kappa\Delta t}{2})}{(1+\frac{\kappa\Delta t}{2})} M(\zeta^n) - \frac{\Delta t}{(1+\frac{\kappa\Delta t}{2})}\left\{\frac{3}{2}Q_*^n - \frac{1}{2}Q_*^{n-1}\right\} = D^n \quad (2.49)$$

where (ζ) is the matrix with (i,j) element ζ_{ij}

M is the Kronecker's product $W^{(x)} \otimes W^{(y)}$, i.e.,

a matrix with (p,q) (i,j) element $W^{(x)}_{pi} W^{(y)}_{qj}$

$W^{(x)}$ is the matrix with (i,j) element $\frac{1}{h}\int \phi_i(x)\phi_j(x)\,dx$

$W^{(y)}$ is the matrix with (i,j) element $\frac{1}{h}\int \phi_i(y)\phi_j(y)\,dy$

$W^{(x)}$ and $W^{(y)}$ are tridiagonal with the local form

$$\frac{1}{6}(1\ 4\ 1)$$

$$Q_*^n = \frac{\varepsilon}{4h^2} J^*(\psi^n, \zeta^n) + \frac{1}{2h}\delta_x(\psi^n) W^{(y)}$$

$\delta_x(\psi)$ is the matrix with (i,j) element $\psi_{i+1,j} - \psi_{i-1,j}$

$J^*(\psi,\zeta)$ is the Arakawa Jacobian which is the finite difference approximation to

$$\frac{1}{3}\left\{[\psi_x\zeta_y - \psi_y\zeta_x] + [(\psi\zeta_y)_x - (\psi\zeta_x)_y + [-(\psi_y\zeta)_x + (\psi_x\zeta)_y]\right\}$$

$$\begin{aligned} J^*_{ij}(\psi,\zeta) = \frac{1}{3}[&(\psi_{i,j-1} + \psi_{i+1,j-1} - \psi_{i,j+1} - \psi_{i+1,j+1})(\zeta_{i+1,j} + \zeta_{ij}) \\ &- (\psi_{i-1,j-1} + \psi_{ij-1} - \psi_{i-1,j+1} - \psi_{ij+1})(\zeta_{ij} + \zeta_{i-1,j}) \\ &+ (\psi_{i+1,j} + \psi_{i+1,j+1} - \psi_{i-1,j} - \psi_{i-1,j+1})(\zeta_{ij+1} + \zeta_{ij}) \\ &- (\psi_{i+1,j-1} + \psi_{i+1,j} - \psi_{i-1,j-1} - \psi_{i-1,j})(\zeta_{ij} + \zeta_{ij-1}) \\ &+ (\psi_{i+1,j} - \psi_{ij+1})(\zeta_{i+1,j+1} + \zeta_{ij}) \\ &- (\psi_{ij-1} - \psi_{i-1,j})(\zeta_{ij} + \zeta_{i-1,j-1}) \\ &+ (\psi_{ij+1} - \psi_{i-1,j})(\zeta_{i-1,j+1} + \zeta_{ij}) \\ &- (\psi_{i+1,j} - \psi_{ij-1})(\zeta_{ij} + \zeta_{i+1,j-1})] \end{aligned} \quad (2.50)$$

The proof of (2.49) can be found in Jespersen (14). This Jacobian is antisymmetric with its arguments and also conserves vorticity, energy and enstrophy when integrated over a closed domain (Arakawa (1)).

Of course, we could have represented (ζ_{ij}) as a vector instead of a matrix, thus eliminating the need for the Kronecker product. The reason for representing (ζ_{ij}) as a matrix is that we can invert the matrix M faster. To see this, we write (2.49) in component form

$$\sum_{i,j} W_{pi}^{(x)} W_{qj}^{(y)} \zeta_{ij}^{n+1} = D_{pq}^{n} \tag{2.51}$$

$$W^{(x)} (\zeta^{n+1}) W^{(y)^t} = D^n \tag{2.52}$$

If (ζ) is an $N \times N$ matrix, we can solve (2.52) by two $N \times N$ matrix inversions--one for $W^{(x)}$ and one $W^{(y)^t}$. Inverting two $N \times N$ matrix is of less complexity than inverting an $N^2 \times N^2$ matrix, which would be required if (ζ) is represented as $N^2 \times 1$ vector. This is one of the benefits gained by using product elements in a rectangular region.

Fix (9) has shown that linear elements for the linearized advective equation $\zeta_t + U\zeta_x = 0$ produce fourth order accurate phase errors. To maintain this accuracy for the vorticity equation, the solution of the Poisson equation (2.41) for the stream function must also be of fourth-order accuracy. This is accomplished by the method of deferred correction (Pereyra (17)).

Note that

$$K(\psi) = h^2 \nabla^2 \psi + \frac{h^4}{12}\left(\nabla^4 - 2\frac{\partial^4}{\partial x^2 \partial y^2}\right)\psi + O(h^6) \tag{2.53}$$

where K is the usual five-point Laplacian

$$K = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix} \tag{2.54}$$

Therefore, two successive Poisson solutions yield ψ to fourth-order accuracy in the following manner. First, obtain a second-order accurate solution, ψ_1, from $K(\psi_1) = h^2\zeta$. Then, a fourth-order accurate estimate of ψ is the solution to

$$K(\psi) = h^2\zeta + \frac{h^4}{12}\left(\nabla^2\zeta - 2\frac{\partial^4}{\partial x^2 \partial y^2}\psi_1\right) \tag{2.55}$$

The boundary condition used in the finite element model is the Charney, Fjörtoft and von Neumann boundary condition, i.e., the specification of streamfunction everywhere on the boundary and of vorticity only at inflow boundary points. To see how the finite element model implements this boundary condition, let us introduce three types of points and their respective computational molecules m, that is, their local contribution to the mass matrix M.

(i) interior points

$$m_I = \frac{1}{36}\begin{pmatrix} 1 & 4 & 1 \\ 4 & 16 & 4 \\ 1 & 4 & 1 \end{pmatrix} \tag{2.56}$$

(ii) regular boundary points (Eastern wall)

$$m_E = \frac{1}{36}\begin{pmatrix} 0 & 1 & 2 \\ 0 & 4 & 8 \\ 0 & 1 & 2 \end{pmatrix} \tag{2.57}$$

(iii) corner points (Southeast corner)

$$m_{SE} = \frac{1}{36}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & \underline{4} \end{pmatrix} \qquad (2.58)$$

The lattice point associated with the given element is underlined. Analogous computational molecules exist for regular boundary points on the northern, southern and western walls and for the southwest, northwest and northeast corner points.

Let

$M_{1/2} = m_I$ for points not at or adjacent to the boundary

$= m_I - \frac{1}{2} m_E$ for interior points adjacent to the eastern boundary but not near a corner

$= m_I - \frac{1}{2} m_E - \frac{1}{2} m_S + \frac{1}{4} m_{SE}$ for the interior southeast corner point

and so on for points adjacent to other boundaries and corners. It can be shown that this new formulation decouples the determination of the interior vorticity from that of the boundary vorticity (thus simplifies the treatment of boundary condition) and is equivalent to

$$M_{1/2}(\zeta^{n+1}) = D^n_{1/2} \qquad (2.59)$$

where $D^n_{1/2}$ has the same relationship to D^n as $M_{1/2}$ to M. In addition

$$M_{1/2} = W^{(x)}_{1/2} \otimes W^{(y)}_{1/2} \qquad \text{where}$$

$$W_{1/2} = \frac{1}{6}\begin{bmatrix} 1/2 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 4 & 1 & \cdots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & 1 & 4 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 1/2 \end{bmatrix} \qquad (2.60)$$

and all of the unknowns are interior points. That is, we use the known dynamic relations between boundary and interior points to disconnect the solution of one from the other.

Furthermore, given the interior values of vorticity from the inversion of M_{η_2}, each of the four boundaries can be decoupled from its neighbors by using an identical strategy. That is, using cornerpoint dynamic relations, the cornerpoint values of vorticity can be eliminated from the solution of the remaining boundary points. For instance, the resulting matrix equation for the eastern boundary minus its cornerpoints can be written

$$\begin{bmatrix} 7 & 2 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 2 & 8 & 2 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 2 & 8 & 2 & \cdots & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdots & 2 & 8 & 2 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 2 & 7 \end{bmatrix} \zeta_E = \mathcal{R}_E \qquad (2.61)$$

where ζ_E is the column vector of unknown vorticity values along the eastern boundary (minus corners) and vector $\mathcal{R}_E$ contains only the known information from previous time steps and the terms reflecting the newly updated values of interior vorticity. When values of vorticity are to be imposed at a particular boundary point, the appropriate row in matrix equation (2.61) and element in $\mathcal{R}_E$ --the i-th, say--are replaced, respectively, by the i-th row of the identity matrix and the desired value of vorticity. The boundary values on the other boundaries (minus cornerpoints) are treated analogously. The cornerpoint values then follow algebraically.

Summing up, the order of calculation in the finite element model is the following: interior vorticity, vorticity at regular boundary points, vorticity at cornerpoints, and lastly, the stream-function. Figure 2.1 is the logic diagram showing the order of the dynamical calculation.

II.2 The Simulation Environment

In this section, we discuss the basic settings of the simulation experiments performed and reported in this thesis. An exterior calculation is performed with the dynamical model in a 1000 km sq. domain with 65x65 grid points. The initial and boundary data for the exterior calculation are supplied by the two barotropic Rossby waves fit to the MODE-I 1500 meter data (McWilliams (16)). The best-fit waves have wavelengths of 171 and 291 kms and periods of 161 and 129 days, respectively. The phase velocities for these waves have components of (-1.0,0.7) and (1.6,2.1) cm/sec. The corresponding group velocities are (0.5, 1.4) and -1.2, 4.2) cm/sec.

The scalings are chosen to be

$$d = 27.25 \text{ km} \tag{2.62}$$

$$V_0 = 1.9 \text{ km/day} \tag{2.63}$$

$$\beta d = 21.2 \text{ days} \tag{2.64}$$

The β-Rossby number is $\varepsilon = \frac{V_0}{\beta d^2} = 1.48$. The bottom frictional Ekman number is chosen to be 0.016 (corresponding to a spin-down time of 1000 days).

The exterior calculation is carried out for six periods (where time is measured in terms of a period of the higher fre-

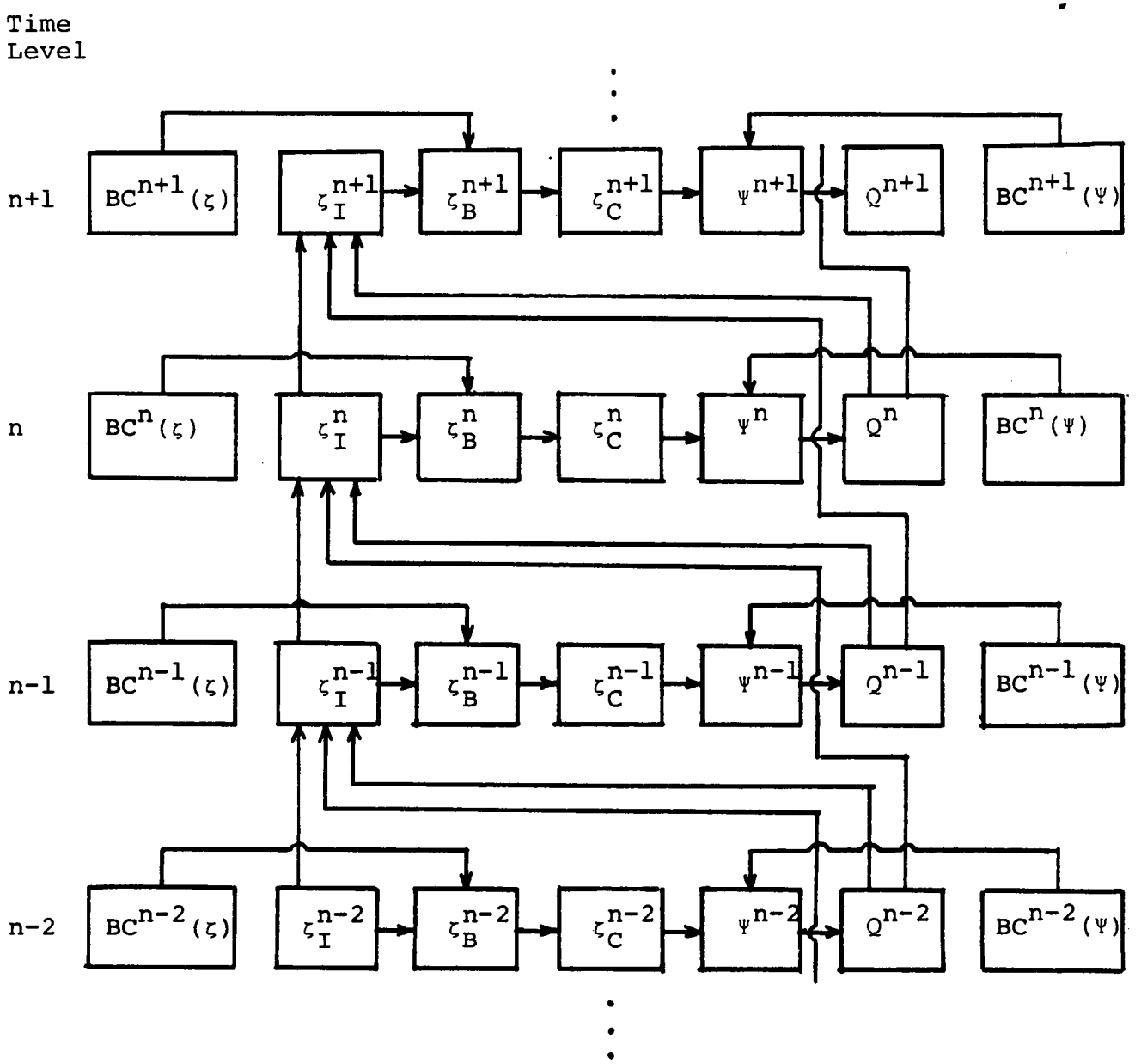


Figure 2.1 The logic diagram showing the order of the dynamical calculation.

quency forcing wave). After two periods, a nonlinear mixture of scales has been achieved within the 1000 km square domain. The values of ψ and ζ, starting at period two, are collected at each time step (there are 128 time steps in one period) to provide initial, boundary and verification data for a forecast calculation over a smaller interior region (500 km square) later. The objective of the exterior calculation is to provide a simulated mid-ocean data base with a plausible statistical character, one which allows controlled forecast simulation experiments under approximate mid-ocean conditions (Robinson (26)).

II. Well-posedness of Open Boundary Conditions

Consider the linearly damped barotropic vorticity equation (2.40)(2.41) in an open domain Ω and a time interval [0,T] (see Figure 2.2). The Charney, Fjörtoft, and von Neumann boundary condition, i.e., the specifications of ψ on Ω_0 and $\partial\Omega\times[0,T]$ and ζ on parts of $\partial\Omega\times[0,T]$ such that $\frac{\partial\psi}{\partial\sigma}>0$ (where σ is the distance measured along $\partial\Omega$ in the counterclockwise direction), is an open boundary condition to the equation. A question arises as to whether this is a well-posed boundary condition, i.e., whether the solution exists for all specifications of the boundary variables involved, is also unique for a particular specification and as well depends continuously on the functions specified. This question turns out to be less than easy.

In the closed domain case, the boundary variables are usually

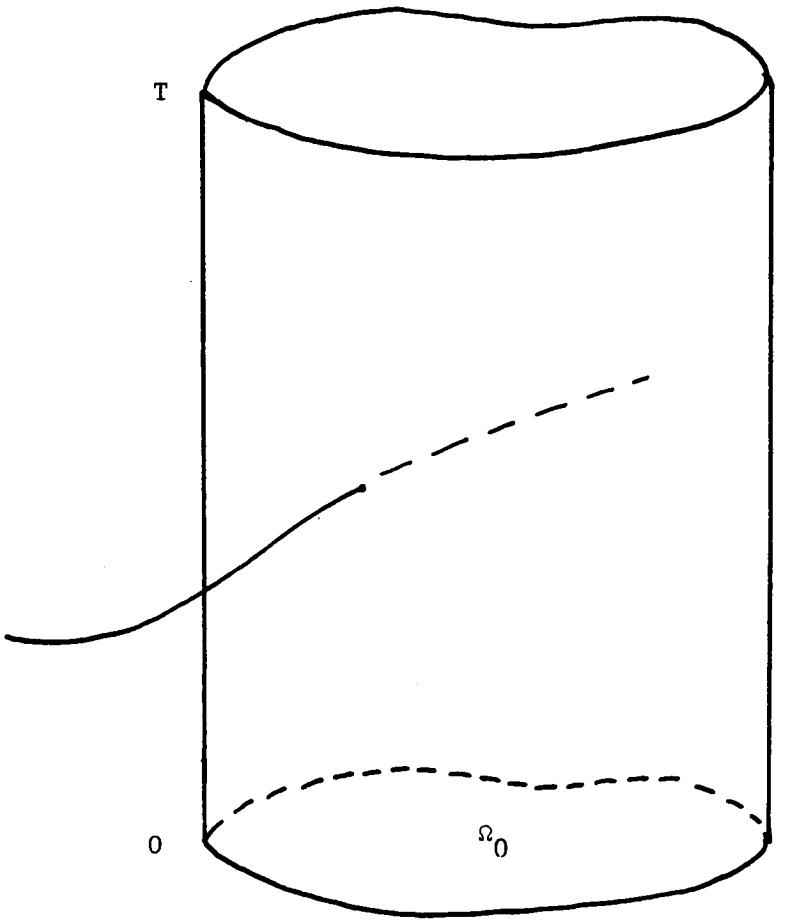


Figure 2.2 An open domain $\Omega * [\ 0, T\ ]$ for the linearly damped barotropic vorticity equation.

fixed to some particular functions, say $\psi = 0$ or $\partial\Omega \times [0,T]$. The only question we have to answer is whether the solution exists and is unique with respect to these particular functions imposed on the boundary. In an open domain case, the situation is slightly different. The variable on the boundary is not restricted to be a fixed function; it can take any of several functions. Let us group all those functions in a functional class such that if the boundary variable takes any of them, the solution to the equation exists. Then what characteristics does this functional class have? Is it a linear space or some sort of Sobolev space? If there are more than two boundary variables involved, are the functional classes of the boundary variables independent? (Here independence means that the choice of a function from the functional class for one boundary variable does not depend on the choice of a function from the functional class for another boundary variable.) It turns out that these classes of functions are very difficult to characterize, and usually are neither linear spaces nor independent.

Take the CFVN boundary condition to the linearly damped barotropic vorticity equation (in nondimensional form) as an example.

$$\frac{\partial \zeta}{\partial t} + \varepsilon J(\psi, \zeta) + \psi_x = -\kappa \zeta \qquad (2.40)$$

$$\nabla^2 \psi = \zeta \qquad (2.41)$$

The CFVN boundary condition is

$$\psi|_{\Omega_0} = \psi_0 \tag{2.65}$$

$$\psi|_{\partial\Omega\times[0,T]} = \psi_B \tag{2.66}$$

$$\zeta|_{\substack{\partial\Omega\times[0,T] \\ \frac{\partial\psi}{\partial\sigma}>0}} = \zeta_B \tag{2.67}$$

Let C_1 be the class of functions ψ_0 is going to take,

C_2 be the class of functions ψ_B is going to take,

C_3 be the class of functions ζ_B is going to take.

If the solution is required to be smooth, then

$$\psi_0|_{\partial\Omega_0} = \psi_B|_{\partial\Omega_0} \tag{2.68}$$

$$\nabla^2\psi_0|_{\substack{\partial\Omega_0 \\ \frac{\partial\psi}{\partial\sigma}>0}} = \zeta_B|_{\substack{\partial\Omega_0 \\ \frac{\partial\psi}{\partial\sigma}>0}} \tag{2.69}$$

$$\frac{\partial\zeta_B}{\partial t} + \varepsilon J(\psi_0, \nabla^2\psi_0) + \psi_{0x} + k\nabla^2\psi_0|_{\substack{\partial\Omega_0 \\ \frac{\partial\psi}{\partial\sigma}>0}} = 0 \tag{2.70}$$

Rewrite (2.40) as

$$\frac{\partial\zeta}{\partial t} + \varepsilon\left[\frac{\partial\psi}{\partial n}\frac{\partial}{\partial\sigma}(\zeta+\frac{y}{\varepsilon}) - \frac{\partial\psi}{\partial\sigma}\frac{\partial}{\partial n}(\zeta+\frac{y}{\varepsilon})\right] = -k\zeta \tag{2.71}$$

At the boundary points where the flow is tangential

$(u_n = -\frac{\partial\psi}{\partial\sigma} = 0)$

$$\frac{\partial\zeta_B}{\partial t} + \varepsilon\left[\frac{\partial\psi}{\partial n}\frac{\partial}{\partial\sigma}(\zeta_B+\frac{y}{\varepsilon})\right] + k\zeta_B|_{\substack{\partial\Omega \\ \frac{\partial\psi}{\partial\sigma}=0}} = 0 \tag{2.72}$$

This is one way to compute $\frac{\partial\psi}{\partial n}$.

$\frac{\partial\psi}{\partial n}|_{\substack{\partial\Omega \\ \frac{\partial\psi}{\partial\sigma}=0}}$ in (2.72) can also be computed from ψ_0, ψ_B, ζ_B at previous times. These two values must coincide.

Bennett (5) concludes from (2.72) that the CFVN boundary

condition is ill-posed because $\frac{\partial \psi}{\partial n}$, which is equal to the tangential velocity, is a part of the solution to be found. This viewpoint seems to be very narrow. If his viewpoint is adopted, then almost all nonlinear initial-boundary value problems are ill-posed, no matter what kind of boundary condition is used.

From (2.68), (2.69), (2.70), and (2.72), we know that C_1, C_2, and C_3 are not independent. They are the constraints the initial and boundary data have to satisfy for the solution to exist. They are clues to the discovery of <u>minimal</u> and <u>nonredundant</u> information required to determine the solution.

To see whether the solution depends continuously on the initial and boundary data, let ψ_1 , ψ_2 be two solutions of equations (2.40), (2.41), $\psi' = \psi_2 - \psi_1$, then

$$\frac{\partial}{\partial t}\nabla^2\psi' + \varepsilon J(\psi_1+\psi', \nabla^2\psi') + \varepsilon J(\psi', \nabla^2\psi_1 + \frac{y}{\varepsilon}) + k\nabla^2\psi' = 0 \qquad (2.73)$$

$$\begin{aligned}
\frac{1}{2}\frac{\partial}{\partial t}\int_\Omega |\nabla^2\psi'|^2 dA &= \int_\Omega (\nabla^2\psi')\frac{\partial}{\partial t}(\nabla^2\psi')dA \\
&= -\int_\Omega (\nabla^2\psi')\left[\varepsilon J(\psi_1+\psi', \nabla^2\psi') + \varepsilon J(\psi', \nabla^2\psi_1 + \frac{y}{\varepsilon}) + k\nabla^2\psi'\right]dA \\
&= \frac{\varepsilon}{2}\oint_{\partial\Omega}(\nabla^2\psi')^2\frac{\partial}{\partial\sigma}(\psi_1+\psi')d\ell \\
&\quad + \varepsilon\int_\Omega \nabla^2\psi' J(\nabla^2\psi_1 + \frac{y}{\varepsilon}, \psi')dA \\
&\quad - k\int_\Omega (\nabla^2\psi')^2 dA
\end{aligned} \qquad (2.74)$$

where the following identities have been used

$$J(\alpha,\beta) = -J(\beta,\alpha) \tag{2.75}$$

$$\int_\Omega \gamma J(\alpha,\beta)dA + \int_\Omega \beta J(\alpha,\gamma)dA = -\oint_{\partial\Omega} \beta\gamma \frac{\partial\alpha}{\partial\sigma} d\ell \tag{2.76}$$

$$\int \beta J(\alpha,\beta)dA = -\frac{1}{2}\oint_{\partial\Omega} \beta^2 \frac{\partial\alpha}{\partial\sigma} d\ell \tag{2.76a}$$

Using the following identity

$$\int_\Omega (\nabla\alpha\cdot\nabla\beta)dA + \int_\Omega (\alpha\nabla^2\beta)dA = \oint_{\partial\Omega} \alpha\frac{\partial\beta}{\partial n} d\ell \tag{2.77}$$

we get

$$\frac{1}{2}\frac{\partial}{\partial t}\int_\Omega \|\nabla\psi'\|^2 dA = \frac{1}{2}\frac{\partial}{\partial t}\int_\Omega (\nabla\psi'\cdot\nabla\psi')dA$$

$$= \int_\Omega (\nabla\psi'\cdot\nabla\frac{\partial\psi'}{\partial t})dA$$

$$= \oint_{\partial\Omega} \psi'\frac{\partial^2\psi'}{\partial n\partial t} d\ell - \int_\Omega (\psi'\frac{\partial}{\partial t}\nabla^2\psi')dA$$

$$= \oint_{\partial\Omega} \psi'\frac{\partial^2\psi'}{\partial n\partial t} d\ell + \int_\Omega \psi'[\varepsilon J(\psi_1+\psi',\nabla^2\psi') + \varepsilon J(\psi',\nabla^2\psi_1+\frac{y}{\varepsilon}) + k^2\nabla^2\psi']dA$$

$$= \oint_{\partial\Omega} [\psi'\frac{\partial^2\psi'}{\partial n\partial t} + \frac{\varepsilon}{2}\psi'^2\frac{\partial}{\partial\sigma}(\nabla^2\psi_1+\frac{y}{\varepsilon}) - \varepsilon\psi'\nabla^2\psi'\frac{\partial}{\partial\sigma}(\psi_1+\psi') + k\psi'\frac{\partial\psi'}{\partial n}]d\ell$$

$$- \varepsilon\int_\Omega \nabla^2\psi' J(\psi_1,\psi')dA - k\int_\Omega \|\nabla\psi'\|^2 dA \tag{2.78}$$

We make the following estimates

$$\int_\Omega \nabla^2\psi' J(\nabla^2\psi_1+\tfrac{y}{\varepsilon},\psi')\,dA \le \int_\Omega |\nabla^2\psi'|\,|J(\nabla^2\psi_1+\tfrac{y}{\varepsilon},\psi')|\,dA$$

$$\le \int_\Omega |\nabla^2\psi'|\,\|\nabla(\nabla^2\psi_1+\tfrac{y}{\varepsilon})\|\,\|\nabla\psi'\|\,dA$$

$$\le \max_\Omega \|\nabla(\nabla^2\psi_1+\tfrac{y}{\varepsilon})\| \left(\frac{1}{2a}\int_\Omega |\nabla^2\psi'|^2 dA + \frac{a}{2}\int_\Omega \|\nabla\psi'\|^2 dA\right) \tag{2.79}$$

$$\int_\Omega \nabla^2\psi' J(\psi_1,\psi')\,dA \le \int_\Omega |\nabla^2\psi'|\,|J(\psi_1,\psi')|\,dA$$

$$\le \int_\Omega |\nabla^2\psi'|\,\|\nabla\psi_1\|\,\|\nabla\psi'\|\,dA$$

$$\le \max_\Omega \|\nabla\psi_1\| \left(\frac{1}{2b}\int_\Omega |\nabla^2\psi'|^2 dA + \frac{b}{2}\int_\Omega \|\nabla\psi'\|^2 dA\right) \tag{2.80}$$

where $$\|\nabla\alpha\| = \left[\left(\frac{\partial\alpha}{\partial x}\right)^2+\left(\frac{\partial\alpha}{\partial y}\right)^2\right]^{1/2} \tag{2.81}$$

a, b are arbitrary positive constants.

Using the estimates (2.79), (2.80) in (2.74), (2.78) respectively, we get

$$\frac{1}{2}\frac{\partial}{\partial t}\int_\Omega |\nabla^2\psi'|^2 dA + k\int_\Omega |\nabla^2\psi'|^2 dA \tag{2.82}$$

$$\le I_1 + \frac{\varepsilon}{2}\max_\Omega \|\nabla(\nabla^2\psi_1+\tfrac{y}{\varepsilon})\| \left(\frac{1}{a}\int_\Omega |\nabla^2\psi'|^2 dA + a\int_\Omega \|\nabla\psi'\|^2 dA\right)$$

$$\frac{1}{2}\frac{\partial}{\partial t}\int_\Omega \|\nabla\psi'\|^2 dA + k\int_\Omega \|\nabla\psi'\|^2 dA$$

$$\le I_2 + \frac{\varepsilon}{2}\max_\Omega \|\nabla\psi_1\| \left(\frac{1}{b}\int_\Omega |\nabla^2\psi'|^2 dA + b\int_\Omega \|\nabla\psi'\|^2 dA\right) \tag{2.83}$$

where
$$I_1 = \frac{\varepsilon}{2}\oint_{\partial\Omega} (\nabla^2\psi')^2 \frac{\partial}{\partial\sigma}(\psi_1+\psi')\,d\ell \tag{2.84}$$

$$I_2 = \oint_{\partial\Omega} \Big[\psi'\frac{\partial^2\psi'}{\partial n\partial t} + \frac{\varepsilon}{2}\psi'^2\frac{\partial}{\partial\sigma}\Big(\nabla^2\psi_1+\frac{y}{\varepsilon}\Big) - \varepsilon\psi'\nabla^2\psi'\frac{\partial}{\partial\sigma}(\psi_1+\psi') + k\psi'\frac{\partial\psi'}{\partial n}\Big]\,d\ell \tag{2.85}$$

Let
$$k_1^2 = \max_{0\le t\le T}\max_{\Omega} \frac{\varepsilon}{2}\|\nabla(\nabla^2\psi_1+\frac{y}{\varepsilon})\| \tag{2.86}$$

$$k_2^2 = \max_{0\le t\le T}\max_{\Omega} \frac{\varepsilon}{2}\|\nabla\psi_1\| \tag{2.87}$$

$$a = b = \frac{k_1}{k_2} \tag{2.88}$$

Multiply (2.82), (2.83) by K_2^2, K_1^2 respectively, and add them together

$$\frac{1}{2}\frac{\partial}{\partial t}\Big(k_2^2\int_\Omega |\nabla^2\psi'|^2 dA + k_1^2\int_\Omega \|\nabla\psi'\|^2 dA\Big)$$
$$\le (k_2^2 I_1 + k_1^2 I_2) + (k_1k_2 - k)\Big(k_2^2\int_\Omega |\nabla^2\psi'|^2 dA + k_1^2\int_\Omega \|\nabla\psi'\|^2 dA\Big) \tag{2.89}$$

Let
$$M = \max_{0\le t\le T}\max_{\Omega}\Big\{\Big|\varepsilon\frac{\partial\psi_2}{\partial\sigma}\Big|, \Big|\frac{\partial^2\psi_1}{\partial n\partial t}\Big|, \Big|\frac{\partial^2\psi_2}{\partial n\partial t}\Big|, \Big|\varepsilon\frac{\partial}{\partial\sigma}\Big(\nabla^2\psi_1+\frac{y}{\varepsilon}\Big)\Big|, \Big|k\frac{\partial\psi_1}{\partial n}\Big|, \Big|k\frac{\partial\psi_2}{\partial n}\Big|\Big\} \tag{2.90}$$

$$I_1 \le \frac{M}{2}\oint_{\partial\Omega} |\nabla^2\psi'|^2 d\ell \tag{2.91}$$

$$I_2 \le 2M\oint_{\partial\Omega}|\psi'|\,d\ell + \frac{M}{2}\oint_{\partial\Omega}|\psi'|^2 d\ell + \frac{M}{2}\oint_{\partial\Omega}|\psi'|^2 d\ell + \frac{M}{2}\oint_{\partial\Omega}|\nabla^2\psi'|^2 d\ell + 2M\oint_{\partial\Omega}|\psi'|\,d\ell \tag{2.92}$$

$$k_2^2 I_1 + k_1^2 I_2 \le M_1 \oint_{\partial\Omega} |\nabla^2\psi'|^2 d\ell + M_2 \oint_{\partial\Omega} |\psi'| d\ell + M_3 \oint_{\partial\Omega} |\psi'|^2 d\ell \qquad (2.93)$$

where

$$M_1 = \left(\frac{k_1+k_2}{\alpha}\right) M \qquad (2.94)$$

$$M_2 = 4k_1^2 M \qquad (2.95)$$

$$M_3 = k_1^2 M \qquad (2.96)$$

Integrate (2.89) from t = 0 to t'

$$\left(k_2^2 \int_\Omega |\nabla^2\psi'|^2 dA + k_1^2 \int_\Omega \|\nabla\psi'\|^2 dA\right)\Big|_{t'}$$

$$\le e^{2(k_1k_2-k)t'}\left(k_2^2 \int_\Omega |\nabla^2\psi'|^2 dA + k_1^2 \int_\Omega \|\nabla\psi'\|^2 dA\right)\Big|_0$$

$$+ 2\, e^{2(k_1k_2-k)t'} \Big\{ M_1 \int_0^{t'} \oint_{\partial\Omega} |\nabla^2\psi'|^2 e^{-2(k_1k_2-k)t} d\ell\, dt$$

$$+ M_2 \int_0^{t'} \oint_{\partial\Omega} |\psi'|\, e^{-2(k_1k_2-k)t} d\ell\, dt$$

$$+ M_3 \int_0^{t'} \oint_{\partial\Omega} |\psi'|^2 e^{-2(k_1k_2-k)t} d\ell\, dt \Big\} \qquad (2.97a)$$

$$\le e^{2(k_1k_2-k)t'} \Big\{ \left(k_2^2 \int_\Omega |\nabla^2\psi'|^2 dA + k_1^2 \int_\Omega \|\nabla\psi'\|^2 dA\right)\Big|_0$$

$$+ 2M_1 \int_0^T \oint_{\partial\Omega} |\nabla^2\psi'|^2 d\ell\, dt + 2M_2 \int_0^T \oint_{\partial\Omega} |\psi'|\, d\ell\, dt$$

$$+ 2M_3 \int_0^T \oint_{\partial\Omega} |\psi'|^2 d\ell\, dt \Big\} \qquad (2.97b)$$

Integrate (2.97b) from t' = 0 to T, we get

$$\int_0^T \left(k_2^2 \int_\Omega |\nabla^2\psi'|^2 dA + k_1^2 \int_\Omega \|\nabla\psi'\|^2 dA\right)\Big|_{t'} dt'$$

$$\le \frac{e^{2(k_1k_2-k)T} - 1}{2(k_1k_2-k)} \Big\{ \left(k_2^2 \int_\Omega |\nabla^2\psi'|^2 dA + k_1^2 \int_\Omega \|\nabla\psi'\|^2 dA\right)\Big|_0$$

$$+ 2M_1 \int_0^T \oint_{\partial\Omega} |\nabla^2\psi'|^2 d\ell\, dt + 2M_2 \int_0^T \oint_{\partial\Omega} |\psi'|\, d\ell\, dt \qquad (2.98)$$

$$+ 2M_3 \int_0^T \oint_{\partial\Omega} |\psi'|^2 d\ell\, dt \Big\}$$

If

$$\int_\Omega |\nabla^2 \psi_0'|^2 dA \to 0 \tag{2.99a}$$

$$\int_\Omega \|\nabla \psi_0'\|^2 dA \to 0 \tag{2.99b}$$

$$\int_0^T \oint_{\partial\Omega} |\zeta_B'|^2 d\ell\, dt \to 0 \tag{2.99c}$$

$$\int_0^T \oint_{\partial\Omega} |\psi_B'|^2 d\ell\, dt \to 0 \tag{2.99d}$$

$$\int_0^T \oint_{\partial\Omega} |\psi_B'| d\ell\, dt \to 0 \tag{2.99e}$$

then

$$\int_0^T \left(k_2^2 \int_\Omega |\nabla^2 \psi'|^2 dA + k_1^2 \int_\Omega \|\nabla \psi'\|^2 dA \right)\Big|_{t'} dt' \to 0 \tag{2.100}$$

Thus, we have shown that the solution depends continuously on the initial and boundary data (the uniqueness of the solution follows from this) if

$$k_1^2 = \max_{0\le t\le T} \max_{\Omega} \frac{\varepsilon}{2} \left\| \nabla\left(\nabla^2 \psi_1 + \frac{y}{\varepsilon}\right) \right\| < \infty \tag{2.86}$$

$$k_2^2 = \max_{0\le t\le T} \max_{\Omega} \frac{\varepsilon}{2} \|\nabla \psi_1\| < \infty \tag{2.87}$$

$$M = \max_{0\le t\le T} \max_{\Omega} \left\{ \left|\varepsilon \frac{\partial \psi_1}{\partial \sigma}\right|, \left|\frac{\partial^2 \psi_1}{\partial n \partial t}\right|, \left|\frac{\partial^2 \psi_2}{\partial n \partial t}\right|, \left|\varepsilon \frac{\partial}{\partial \sigma}\left(\nabla^2 \psi_1 + \frac{y}{\varepsilon}\right)\right|, \left|k^2 \frac{\partial \psi_1}{\partial n}\right|, \left|k^2 \frac{\partial \psi_2}{\partial n}\right| \right\} \tag{2.90}$$
$$< \infty$$

also the initial and boundary data are restricted to change in a way such that the solution still exists for the new data and

$$\int_\Omega |\nabla^2 \psi_0'|^2 dA < \infty \tag{2.99a}$$

$$\int_\Omega \|\nabla \psi_0'\|^2 dA < \infty \tag{2.99b}$$

$$\int_0^T \oint_{\partial\Omega} |\zeta_B'|^2 \, dl \, dt < \infty \tag{2.99c}$$

$$\int_0^T \oint_{\partial\Omega} |\psi_B'| \, dl \, dt < \infty \tag{2.99d}$$

$$\int_0^T \oint_{\partial\Omega} |\psi_B'|^2 \, dl \, dt < \infty \tag{2.99e}$$

If we put all these constraints on the functional classes C_1, C_2, C_3, then C_1, C_2, C_3 are not linear spaces anymore. These constraints are only <u>sufficient</u> conditions for the initial-boundary value problem to be well-posed. What the <u>necessary</u> and <u>sufficient</u> conditions on C_1, C_2, C_3 are for the problem to be well-posed is still unknown.

The well-posedness of a certain boundary condition for the quasigeostrophic potential vorticity equation can be treated in the same way. The following results are presented without detailed explanation.

Consider the quasigeostrophic potential vorticity equation in a domain $\Omega \times [0, H] \times [0, T]$ shown in Figure 2.3.

$$\frac{\partial q}{\partial t} + J(\psi, q) = 0 \tag{2.30a}$$

$$\nabla_H^2 \psi + \frac{1}{\beta}\left(\frac{\psi_z}{S(z)}\right)_z + \beta y = q \tag{2.32a}$$

where

$$\nabla_H^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \tag{2.33}$$

In the following, we assume $S(z) > 0$ which corresponds to the physical situation that the basic state of the fluid is stably stratified. A plausible boundary condition for (2.30a),

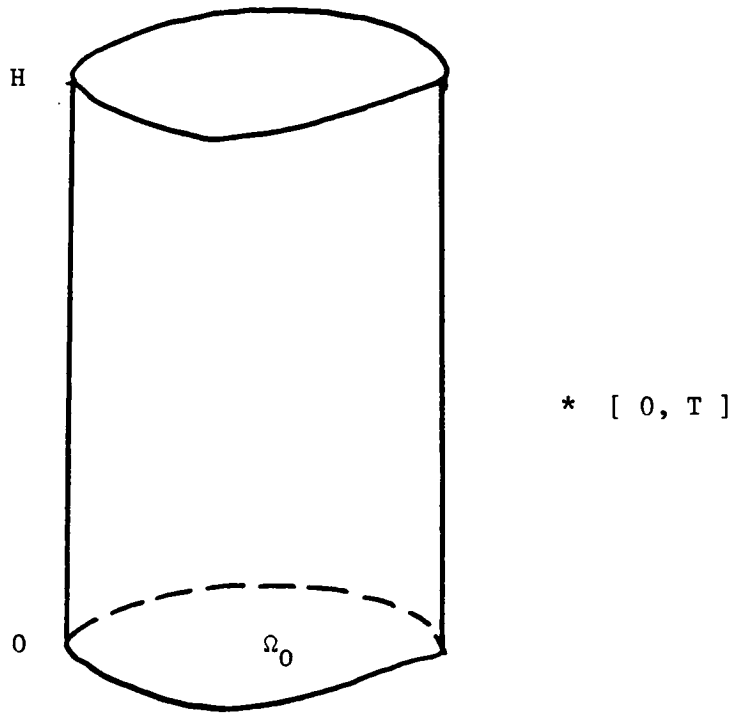


Figure 2.3 An open domain Ω * [0, H] * [0, T] for the quasigeostrophic potential vorticity equation.

(2.32a) is

$$\psi|_{\Omega(t=0)\times[0,H]} = \psi_0 \tag{2.101}$$

$$\psi|_{\partial\Omega\times[0,H]\times[0,T]} = \psi_B \tag{2.102}$$

$$\psi|_{\Omega(z=0)\times[0,T]} = \psi_L \tag{2.103}$$

$$\psi|_{\Omega(z=H)\times[0,T]} = \psi_U \tag{2.104}$$

$$q|_{\partial\Omega\times[0,H]\times[0,T]} = q_B \tag{2.105}$$

$$\frac{\partial\psi}{\partial\sigma} > 0$$

Let C_1 be the class of functions that ψ_0 is going to take,

C_2 be the class of functions that ψ_B is going to take,

C_3 be the class of functions that ψ_L is going to take,

C_4 be the class of functions that ψ_U is going to take,

C_5 be the class of functions that q_B is going to take.

The following constraints must be satisfied for the solution to exist

$$\psi_0|_{\partial\Omega(t=0)\times[0,H]} = \psi_B|_{\partial\Omega(t=0)\times[0,H]} \tag{2.106}$$

$$\psi_0|_{\Omega(t=0,z=0)} = \psi_L|_{\Omega(t=0,z=0)} \tag{2.107}$$

$$\psi_0|_{\Omega(t=0,z=H)} = \psi_U|_{\Omega(t=0,z=H)} \tag{2.108}$$

$$\psi_B|_{\partial\Omega(z=0)\times[0,T]} = \psi_L|_{\partial\Omega(z=0)\times[0,T]} \tag{2.109}$$

$$\psi_B|_{\partial\Omega(z=H)\times[0,T]} = \psi_U|_{\partial\Omega(z=H)\times[0,T]} \tag{2.110}$$

$$q_B\Big|_{\substack{\partial\Omega(t=0)\times[0,H] \\ \frac{\partial\psi}{\partial\sigma}>0}} = \nabla_H^2\psi_0 + \frac{1}{B}\left(\frac{\psi_{0z}}{S}\right)_z + \beta y\Big|_{\substack{\partial\Omega(t=0)\times[0,H] \\ \frac{\partial\psi}{\partial\sigma}>0}} \tag{2.111}$$

$$\frac{\partial q_B}{\partial t} + J\left(\psi_0, \nabla_H^2\psi_0 + \frac{1}{B}\left(\frac{\psi_{0z}}{S}\right)_z + \beta y\right)\Big|_{\substack{\partial\Omega(t=0)\times[0,H] \\ \frac{\partial\psi}{\partial\sigma}>0}} = 0 \tag{2.112}$$

$$\frac{\partial q_B}{\partial t} + \frac{\partial\psi}{\partial n}\frac{\partial q_B}{\partial\sigma}\Big|_{\substack{\partial\Omega\times[0,H] \\ \frac{\partial\psi}{\partial\sigma}=0}} = 0 \tag{2.113}$$

where $\frac{\partial\psi}{\partial n}\Big|_{\substack{\partial\Omega\times[0,H] \\ \frac{\partial\psi}{\partial\sigma}=0}}$ in (2.113) is related to $\psi_0, \psi_B, \psi_L, \psi_U$, q_B at previous times.

Let ψ_1, ψ_2 be two solutions of (2.30a) (2.32a) and $\psi' = \psi_2 - \psi_1$, then

$$\frac{\partial}{\partial t}\left[\nabla_H^2\psi' + \frac{1}{B}\left(\frac{\psi_z'}{S}\right)_z\right] + J\left(\psi_1+\psi', \nabla_H^2\psi' + \frac{1}{B}\left(\frac{\psi_z'}{S}\right)_z\right)$$

$$+ J\left(\psi', \nabla_H^2\psi_1 + \frac{1}{B}\left(\frac{\psi_{1z}}{S}\right)_z + \beta y\right) = 0 \tag{2.114}$$

The following estimate is derived similarly to (2.89)

$$\frac{1}{2}\frac{\partial}{\partial t}\left(k_2^2\int_0^H\int_\Omega\left[\nabla_H^2\psi' + \frac{1}{B}\left(\frac{\psi_z'}{S}\right)_z\right]^2 dA\,dz + k_1^2\int_0^H\int_\Omega\left[\left(\frac{\partial\psi'}{\partial x}\right)^2\right.\right.$$

$$\left.\left. + \frac{1}{BS}\left(\frac{\partial\psi'}{\partial z}\right)^2 dA\,dz\right) \le k_2^2 I_1 + k_1^2 I_2$$

$$+ k_1 k_2\left(k_2^2\int_0^H\int_\Omega\left[\nabla_H^2\psi' + \frac{1}{B}\left(\frac{\psi_z'}{S}\right)_z\right]^2 dA\,dz\right.$$

$$\left. + k_1^2\int_0^H\int_\Omega\left[\left(\frac{\partial\psi'}{\partial x}\right)^2 + \left(\frac{\partial\psi'}{\partial y}\right)^2 + \frac{1}{BS}\left(\frac{\partial\psi'}{\partial z}\right)^2\right] dA\,dz\right.$$

(2.115)

where

$$I_1 = \frac{1}{2}\int_0^H \oint_{\partial\Omega} \left[\nabla_H^2\psi' + \frac{1}{B}\left(\frac{\psi_z'}{S}\right)_z\right]^2 \frac{\partial(\psi_1+\psi')}{\partial\sigma}\, d\ell\, dz \qquad (2.116)$$

$$I_2 = \int_0^H \oint_{\partial\Omega} \Big[\psi \frac{\partial^2\psi'}{\partial n \partial t} + \frac{\psi'^2}{2}\frac{\partial}{\partial\sigma}\left(\nabla_H^2\psi_1 + \frac{1}{B}\left(\frac{\psi_{1z}}{S}\right)_z + \beta y\right)$$

$$- \psi'\left(\nabla_H^2\psi' + \frac{1}{B}\left(\frac{\psi_z'}{S}\right)_z\right)\frac{\partial(\psi_1+\psi')}{\partial\sigma}\Big]\, d\ell\, dz$$

$$+ \int_{\Omega(z=H)} \psi' \frac{\partial^2\psi'}{\partial z \partial t}\left(1 + \frac{1}{BS(H)}\right) dA - \int_{\Omega(z=0)} \psi' \frac{\partial^2\psi'}{\partial z \partial t}\left(1 + \frac{1}{BS(0)}\right) dA \qquad (2.117)$$

$$K_1^2 = \max_{\Omega\times[0,H]\times[0,T]} \left|\nabla_H\left(\nabla_H^2\psi_1 + \frac{1}{B}\left(\frac{\psi_{1z}}{S}\right)_z + \beta y\right)\right| \qquad (2.118)$$

$$K_2^2 = \max_{\Omega\times[0,H]\times[0,T]} \left|\nabla_H\psi_1\right| \qquad (2.119)$$

Let

$$M = \max_{\Omega\times[0,H]\times[0,T]} \left\{\left|\frac{\partial\psi_2}{\partial\sigma}\right|, \left|\frac{\partial^2\psi_1}{\partial n\partial t}\right|, \left|\frac{\partial^2\psi_2}{\partial n\partial t}\right|,\right.$$

$$\left.\left|\frac{\partial}{\partial\sigma}\left(\nabla_H^2\psi_1 + \frac{1}{B}\left(\frac{\psi_{1z}}{S}\right)_z + \beta y\right)\right|, \left|\frac{\partial^2\psi_1}{\partial z\partial t}\left(1 + \frac{1}{BS(z)}\right)\right|\right\} \qquad (2.120)$$

then

$$\Big(K_2^2\int_0^H\int_\Omega \left[\nabla_H^2\psi' + \frac{1}{B}\left(\frac{\psi_z'}{S}\right)_z\right]^2 dA\, dz$$

$$+ K_1^2\int_0^H\int_\Omega \left[\left(\frac{\partial\psi'}{\partial x}\right)^2 + \left(\frac{\partial\psi'}{\partial y}\right)^2 + \frac{1}{BS}\left(\frac{\partial\psi'}{\partial z}\right)^2\right] dA\, dz\Big)\Big|_{t'}$$

$$\le e^{2k_1k_2t'}\Big(K_2^2\int_0^H\int_\Omega \left[\nabla_H^2\psi_0' + \frac{1}{B}\left(\frac{\psi_{0z}'}{S}\right)_z\right]^2 dA\, dz$$

$$+ K_1^2\int_0^H\int_\Omega \left[\left(\frac{\partial\psi_0'}{\partial x}\right)^2 + \left(\frac{\partial\psi_0'}{\partial y}\right)^2 + \frac{1}{BS}\left(\frac{\partial\psi_0'}{\partial z}\right)^2\right] dA\, dz \qquad \text{(cont.)}$$

$$+ 2M_1 \int_0^{t'} \int_0^H \oint_{\partial\Omega} |q_B'|^2 e^{-2k_1k_2t} \, d\ell \, dz \, dt$$
$$+ 2M_2 \int_0^{t'} \int_0^H \oint_{\partial\Omega} |\psi_B'| \, e^{-2k_1k_2t} \, d\ell \, dz \, dt$$
$$+ 2M_3 \int_0^{t'} \int_0^H \oint_{\partial\Omega} |\psi_B'|^2 e^{-2k_1k_2t} \, d\ell \, dz \, dt$$
$$+ 2M_3 \int_0^{t'} \int_{\Omega(z=H)} |\psi_U'| \, e^{-2k_1k_2t} \, dA \, dt + 2M_3 \int_0^{t'} \int_{\Omega(z=0)} |\psi_L'| \, e^{-2k_1k_2t} \, dA \, dt \Big) \tag{2.121}$$

where

$$M_1 = \frac{k_1^2 + k_2^2}{2} M \tag{2.122}$$

$$M_2 = 2k_1^2 M \tag{2.123}$$

$$M_3 = k_1^2 M \tag{2.124}$$

Also

$$k_2^2 \int_0^T \int_0^H \int_\Omega \left[\nabla_H^2 \psi' + \frac{1}{B}\left(\frac{\psi_z'}{S}\right)_z\right]^2 dA \, dz \, dt + k_1^2 \int_0^T \int_0^H \int_\Omega \left[\left(\frac{\partial \psi'}{\partial x}\right)^2 + \left(\frac{\partial \psi'}{\partial y}\right)^2 + \frac{1}{BS}\left(\frac{\partial \psi'}{\partial z}\right)^2\right] dA \, dz \, dt$$
$$\leq \frac{e^{2k_1k_2T} - 1}{2k_1k_2} \Big\{ k_2^2 \int_0^H \int_\Omega \left[\nabla_H^2 \psi_0' + \frac{1}{B}\left(\frac{\psi_{0z}'}{S}\right)_z\right]^2 dA \, dz$$
$$+ k_1^2 \int_0^H \int_\Omega \left[\left(\frac{\partial \psi_0'}{\partial x}\right)^2 + \left(\frac{\partial \psi_0'}{\partial y}\right)^2 + \frac{1}{BS}\left(\frac{\partial \psi_0'}{\partial z}\right)^2\right] dA \, dz$$
$$+ 2M_1 \int_0^T \int_0^H \oint_{\partial\Omega} |q_B'|^2 \, d\ell \, dz \, dt + 2M_2 \int_0^T \int_0^H \oint_{\partial\Omega} |\psi_B'| \, d\ell \, dz \, dt$$
$$+ 2M_3 \int_0^T \int_0^H \oint_{\partial\Omega} |\psi_B'|^2 \, d\ell \, dz \, dt + 2M_3 \int_0^T \int_{\Omega(z=H)} |\psi_U'| \, dA \, dt$$
$$+ 2M_3 \int_0^T \int_{\Omega(z=0)} |\psi_L'| \, dA \, dt \Big\} \tag{2.125}$$

Thus, if

$$\int_0^H \int_\Omega \left[\nabla_H^2 \psi_0' + \frac{1}{B}\left(\frac{\psi_{0z}'}{S}\right)_z\right]^2 dA \, dz \to 0 \tag{2.126a}$$

$$\int_0^H \int_\Omega \left[\left(\frac{\partial \psi_0'}{\partial x}\right)^2 + \left(\frac{\partial \psi_0'}{\partial y}\right)^2 + \frac{1}{BS}\left(\frac{\partial \psi_0'}{\partial z}\right)^2 \right] dA\, dz \to 0 \qquad (2.126b)$$

$$\int_0^T \int_0^H \oint_{\partial\Omega} |q_B'|\, d\ell\, dz\, dt \to 0 \qquad (2.126c)$$

$$\int_0^T \int_0^H \oint_{\partial\Omega} |\psi_B'|\, d\ell\, dz\, dt \to 0 \qquad (2.126d)$$

$$\int_0^T \int_0^H \oint_{\partial\Omega} |\psi_B'|^2\, d\ell\, dz\, dt \to 0 \qquad (2.126e)$$

$$\int_0^T \int_{\Omega(z=H)} |\psi_U'|\, dA\, dt \to 0 \qquad (2.126f)$$

$$\int_0^T \int_{\Omega(z=0)} |\psi_L'|\, dA\, dt \to 0 \qquad (2.126g)$$

then

$$K_2^2 \int_0^T \int_0^H \int_\Omega \left[\nabla_H^2 \psi' + \frac{1}{B}\left(\frac{\psi_z'}{S}\right)_z \right]^2 dA\, dz\, dt$$

$$+ K_1^2 \int_0^T \int_0^H \int_\Omega \left[\left(\frac{\partial \psi'}{\partial x}\right)^2 + \left(\frac{\partial \psi'}{\partial y}\right)^2 + \frac{1}{BS}\left(\frac{\partial \psi'}{\partial z}\right)^2 \right] dA\, dz\, dt \to 0 \qquad (2.127)$$

So the solution depends continuously on the initial and boundary data (hence it is unique) if the data can only vary in a way such that the solution exists, and if the additional constraints $K_1^2 < \infty$, $K_2^2 < \infty$, $M < \infty$ and (2.126a-g) on C_1, C_2, C_3, C_4 and C_5 are satisfied. Again these are only sufficient conditions. The necessary and sufficient conditions are yet to be found. An interpretation of the formulas (2.97)(2.125) is in the next section.

II.4 The Zero Initial Condition Problem

When we were studying in the previous section the continuous dependence of a solution on the initial and boundary data, we obtained as a bonus an estimate of the error of the solution at time t′ as a function of the errors of the initial and boundary

data accumulated from time 0 to t′. For example, for the linearly damped barotropic vorticity equation, we obtain

$$
\begin{aligned}
&\left(K_2^2 \int_\Omega |\nabla^2\psi'|^2 dA + K_1^2 \int_\Omega \|\nabla\psi'\|^2 dA \right)\Big|_{t'} \\
&\quad \le e^{2(K_1K_2-k)t'}\left(K_2^2 \int_\Omega |\nabla^2\psi_0'|^2 dA + K_1^2 \int_\Omega \|\nabla\psi_0'\|^2 dA \right) \\
&\quad\quad + 2\, e^{2(K_1K_2-k)t'} \Big\{ M_1 \int_0^{t'} \oint_{\partial\Omega} |\zeta_B'|^2 e^{-2(K_1K_2-k)t} \, d\ell \, dt \\
&\quad\quad\quad + M_2 \int_0^{t'} \oint_{\partial\Omega} |\psi_B'| \, e^{-2(K_1K_2-k)t} \, d\ell \, dt \\
&\quad\quad\quad + M_3 \int_0^{t'} \oint_{\partial\Omega} |\psi_B'|^2 e^{-2(K_1K_2-k)t} \, d\ell \, dt \Big\}
\end{aligned}
\tag{2.97a}
$$

However, this estimate is too conservative. The error is estimated to grow with time only at a rate no faster than the exponential. This is not a very tight bound. In practice, the error grows with time at a rate slower than the exponential, say a polynomial. In some cases it may decrease. For instance, consider a dynamical calculation for which perfect boundary data are available while the initial data are not known and hence taken to be zero. Starting from zero initial conditions, we drive the dynamical model forward in time with perfect data supplied on the boundary for all time. It is reasonable to expect that the error will decrease with time because more information about the state of the system is provided through the boundary data. The more we know about the system, the better the solution is. Therefore, although (2.97a) is of some theoretical interest, it is not practically useful. However, it is very difficult to find a tighter bound without losing the gener-

ality. Here we consider a special case for which a exact error bound can be established.

Consider the linear barotropic vorticity equation with linear damping (i.e., $\varepsilon = 0$ in (2.40) and (2.41))

$$\frac{\partial}{\partial t}\nabla^2\psi + \psi_x = -k\nabla^2\psi \tag{2.128}$$

in an open domain shown in Figure (2.2)

Let ψ_0 be a solution of (2.128) which satisfies the following initial and boundary conditions

$$\psi_0|_{\Omega_0} = f \tag{2.129}$$

$$\psi_0|_{\partial\Omega\times[0,T]} = g \tag{2.130}$$

We want to estimate the error growth rate, starting from a zero initial condition (hence with an initial condition error) but using a perfect boundary condition. Let ψ_1 be the solution to this problem

$$\frac{\partial}{\partial t}\nabla^2\psi_1 + \psi_{1x} = -k\nabla^2\psi_1 \tag{2.131}$$

$$\psi_1|_{\Omega_0} = 0 \tag{2.132}$$

$$\psi_1|_{\partial\Omega\times[0,T]} = g \tag{2.133}$$

and $\psi' = \psi_1 - \psi_0$ then

$$\frac{\partial}{\partial t}\nabla^2\psi' + \psi_x' = -k\nabla^2\psi' \tag{2.134}$$

$$\psi'|_{\Omega_0} = f \tag{2.135}$$

$$\psi'|_{\partial\Omega\times[0,T]} = 0 \tag{2.136}$$

It is easy to show that

$$\frac{1}{2}\frac{\partial}{\partial t}\int_\Omega \|\nabla\psi'\|^2 dA = \oint_{\partial\Omega}\left[\psi'\frac{\partial^2\psi'}{\partial n\partial t} + k\psi'\frac{\partial\psi'}{\partial n} + \frac{1}{2}\psi'^2 n_1\right]dl - k\int_\Omega \|\nabla\psi'\|^2 dA \tag{2.137}$$

Since $\psi' = 0$ on $\partial\Omega \times [0, T]$

$$\frac{1}{2}\frac{\partial}{\partial t}\int_\Omega \|\nabla\psi'\|^2 dA = -k\int_\Omega \|\nabla\psi'\|^2 dA \tag{2.137a}$$

$$\int_\Omega \|\nabla\psi'\|^2 dA\Big|_{t'} = e^{-2kt'}\int_\Omega \|\nabla f\|^2 dA \tag{2.138}$$

Thus, the error decreases exponentially with time. The e-folding time is equal to $\frac{1}{2k}$. If $k = 0$, then

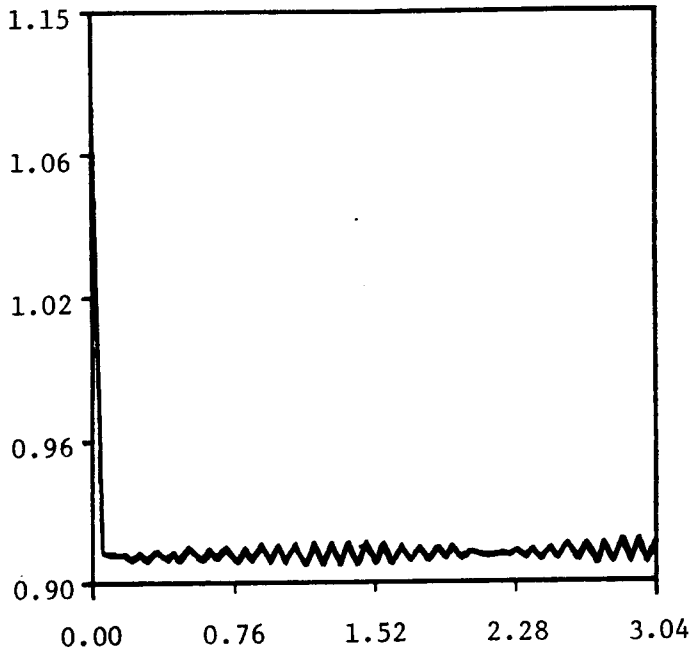
$$\int_\Omega \|\nabla\psi'\|^2 dA\Big|_{t'} = \int_\Omega \|\nabla f\|^2 dA \tag{2.138a}$$

The error neither increases nor decreases with time, but stays constant. Figure 2.4 shows the errors as a function of time from numerical calculations for the case $k = 0$. Figure (2.4a) corresponds to the case with a two Rossby Waves true solution

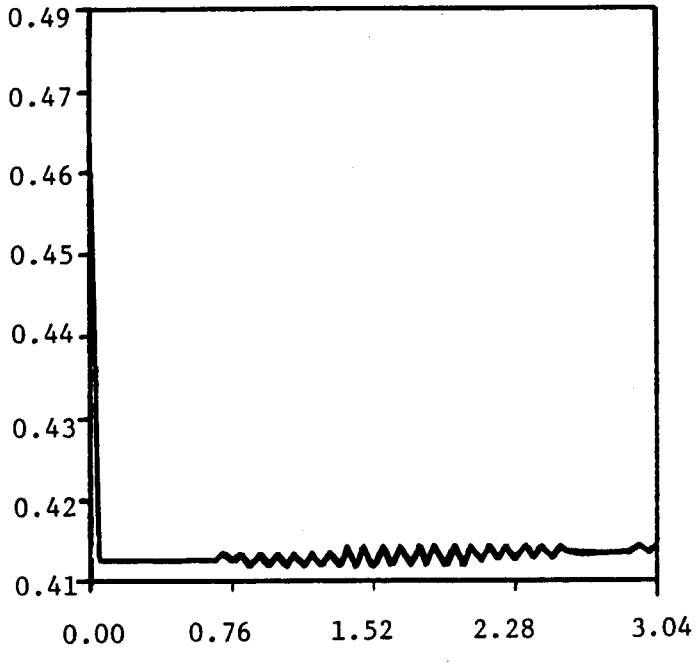
$$\psi_0 = \cos(-0.83|x + 0.559y + 0.828t - 0.55) + 1.93\cos(-0.354x - 0.469y + 1.028t + 2.07).$$

Figure (2.4b) corresponds to the case with a one Rossby Wave true solution $\psi_0 = \cos(-0.83|x + 0.559y + 0.828t - 0.55)$.

And we do see that the errors stay almost constant. To see how the error evolves in time with the initial and boundary condition errors in the general case, we have to resort to numerical methods.



(a)



(b)

Figure 2.4 The energies of the errors as functions of time for the linear experiments (k=0) that start with zero initial conditions corresponding to the true solutions

(a) $\Psi_0 = \cos(-0.831x+0.559y+0.828t-0.55) + 1.93\cos(-0.354x-0.469y+1.025t+2.07)$

(b) $\Psi_0 = \cos(-0.831x+0.559y+0.828t-0.55)$

II.5 Summary of Detailed Persistent Dynamical Forecast Results

In this section, we perform some benchmark dynamical simulation experiments. The purpose of these benchmark calculations is to provide results with which future simulation experiments can be compared. Obviously, the first benchmark dynamical calculation is the one that uses perfect initial and boundary data. (Here, and hereafter, the data are taken from the data set generated by the exterior calculation in Section II.1). The only error sources in this experiment are purely computational--truncation and roundoff errors. It has been performed and was found to generate an error of only about 1 percent after one period of calculation. This is an illustration of the accuracy of our dynamical model.

The next set of benchmark calculations is the one in which we have a perfect initial condition, but not a boundary condition. The boundary condition is thus kept frozen from the initial time. This represents complete ignorance of the boundary condition. The result is to be compared against the result in which we have partial information on the boundary condition. This comparison will give us a measure of the merits of this additional boundary condition and hence give some clue as to the most efficient boundary condition sampling scheme.

The persistent dynamical calculation has been performed for 60 time steps starting from period 2.25, 3, 3.75, and 4.5 respectively. The NRMS (Normalized Root Mean Square) streamfunction error, which is defined as

$$\varepsilon = \left[\frac{\iint (\psi_f - \psi_t)^2 dA}{\iint \psi_t^2 dA} \right]^{1/2} \qquad (2.139)$$

where ψ_t is the true field

ψ_f is the forecast field,

is shown in Figure 2.5. Figure 2.5a is for the experiment that starts from period 2.25. Figure 2.5b-d from period 3, 3.75, 4.5 respectively. The error, starting from zero, increases to 100 percent in about a month. The streamfunction error field is seen to have a steady state gross feature after some transience. This steady state gross feature of the streamfunction error field is shown in Figure 2.6a-d for starting time at period 2.25, 3, 3.75, 4.5 respectively. These error maps are to be used to construct error models in Chapter IV.

II.6 Dependence on Persistence Interval and Reinitialization Interval

In a persistent dynamical forecast, the forecast error increases very rapidly with time. The NRMS streamfunction error reaches 100 per cent in about a month. To get a better forecast, we can update either the boundary condition and/or the interior field. In this section, several dynamical simulation experiments are performed with different persistent intervals and reinitialization intervals. The former refer to how frequently the boundary condition is updated, while the latter refer to how frequently the interior field is updated. A Pn experiment denotes one in which the boundary condition is updated every n time steps but kept persistent between them. A Rn experiment denotes one in which the interior field is updated every n time steps. We are interested in knowing how the

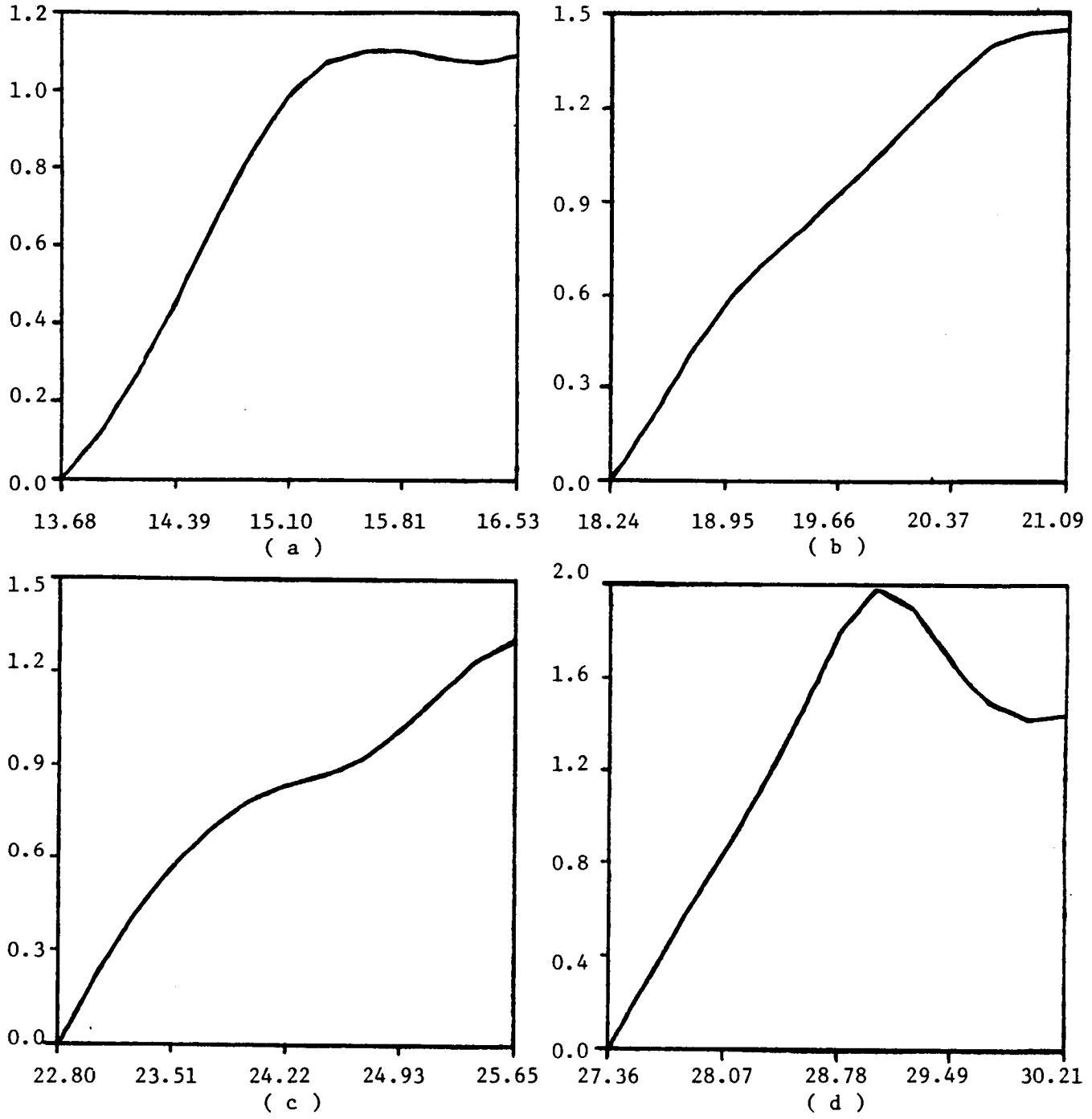
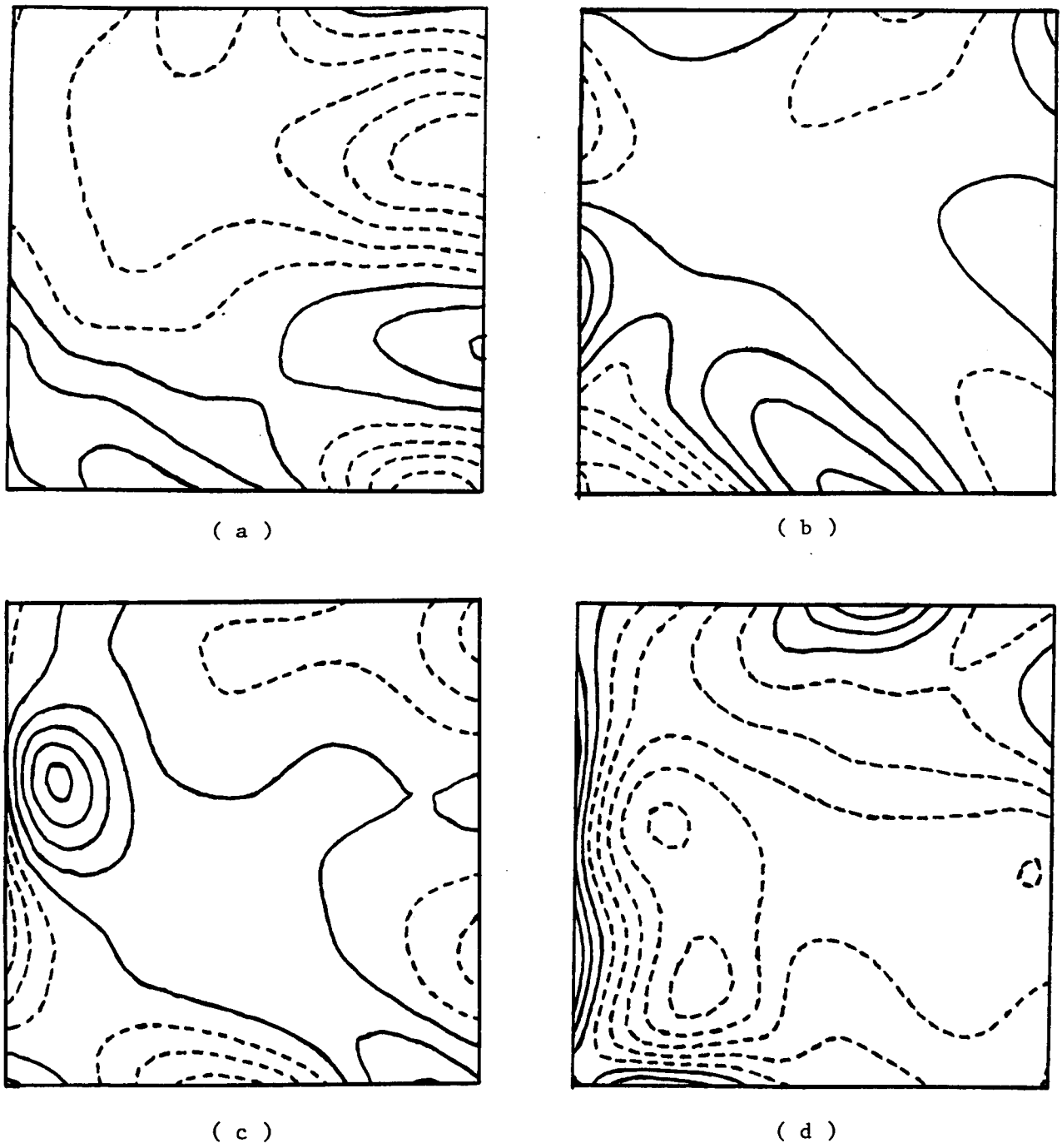


Figure 2.5 NRMS streamfunction errors as functions of time for the persistent dynamical forecasts that start at period
(a) 2.25 (b) 3.00
(c) 3.75 (d) 4.50



(a) (b)

(c) (d)

Figure 2.6 The steady state gross feature of the streamfunction error fields for the persistent dynamical forecasts that start at period
(a) 2.25 (b) 3.00
(c) 3.75 (d) 4.50

forecast error depends on the persistent interval and the re-initialization interval. To determine the dependencies and sensitivities of the forecast error on various parameters constitutes one of our major modeling efforts.

Some results related to boundary updating have been reported in Robinson (26). It was found there that the global RMS streamfunction error is controlled essentially by the time interval between successive updates of individual points on the boundary. It does not depend strongly on whether the boundary points are updated all at once or piecewise in some particular order, although piecewise updating does smooth the error curve. Also, the forecast error, although sensitive to the overall accuracy of the supplied boundary streamfunction data, is relatively insensitive to large errors in boundary values of vorticity. Thus, in the experiments reported here, when we update the boundary condition, we update only the streamfunction and at one time. In the interior updating, we update the streamfunction by true values. The vorticity is updated by the derived vorticity from the streamfunction with a 2nd order Laplacian scheme.

All the experiments performed in this regard are summarized in Table 2.1. They all start from perfect initial conditions but use different boundary and interior updating strategies. Whenever a field is updated in the interior, it is updated at two consecutive time steps in order for it to be fully re-initialized (see Figure 2.1).

Experiment No.	Boundary Updating	Interior Updating
2.1	P10	---
2.2	P10	R5
2.3	P10	R10
2.4	P10	R30
2.5	P15	---
2.6	P15	R5
2.7	P15	R10
2.8	P15	R30

Table 2.1 -- Summary of boundary updating and interior updating experiments.

The NRMS streamfunction errors for these experiments are shown in Figure 2.7 and Figure 2.8. Figure 2.7a-d is for Experiment No. 2.1-4, while Figure 2.8a-d is for Experiment No. 2.5-8. It can be seen that the error is essentially dominated by the boundary condition updating. The interior updating does not do much good if the boundary condition fails to be updated, and is not effective until some time has passed (around 30 days)*. This tells us that, in a forecasting experiment, the most efficient way of collecting data to obtain a better forecast is to collect the data on the boundary.

---

* Note the envelopes of the error curves, especially near the end of the calculations.

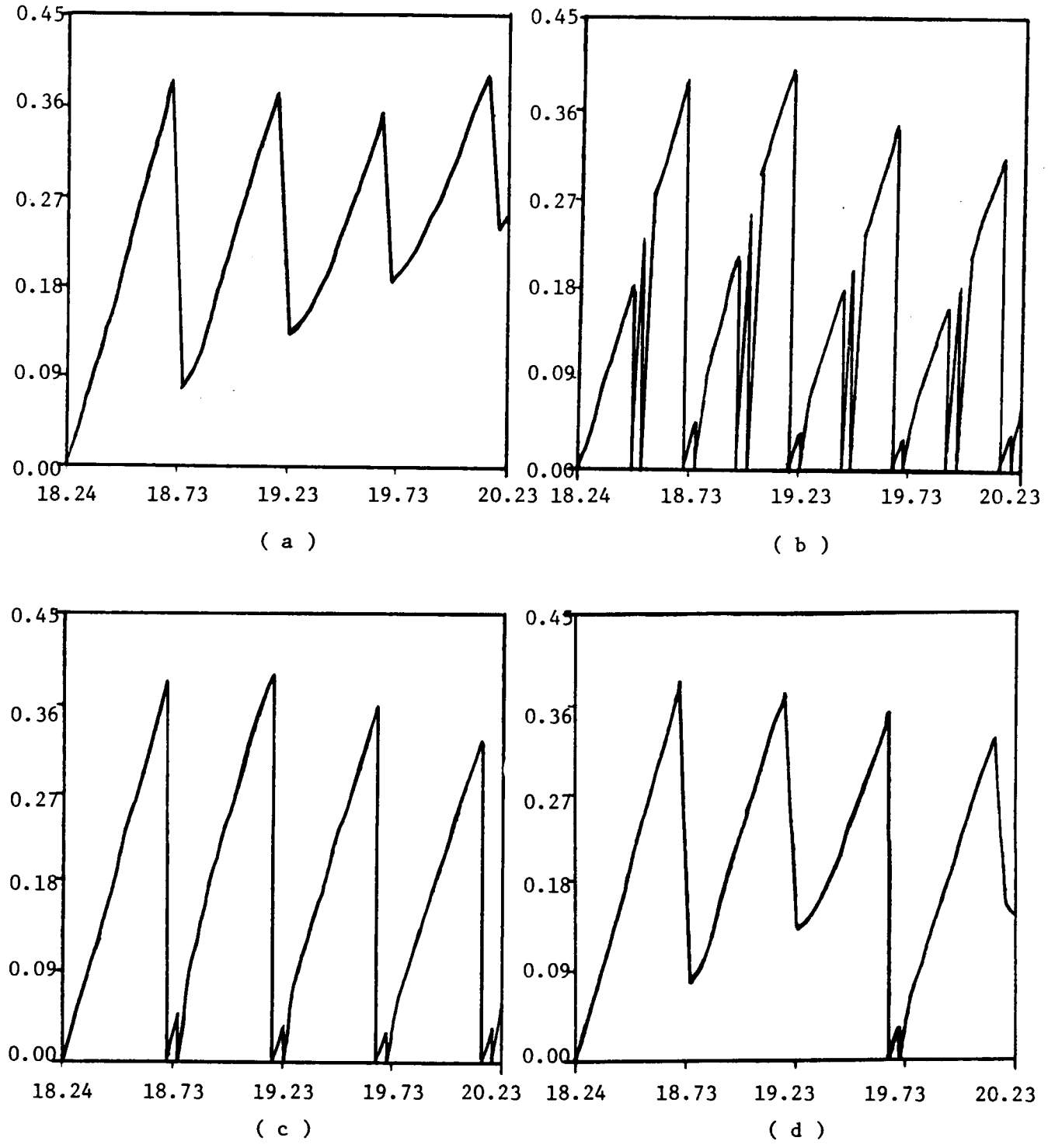


Figure 2.7 NRMS streamfunction errors for the experiments that update the boundary every 10 time steps, but update the interior every
(a) ---------- (b) 5 time steps
(c) 10 time steps (d) 30 time steps

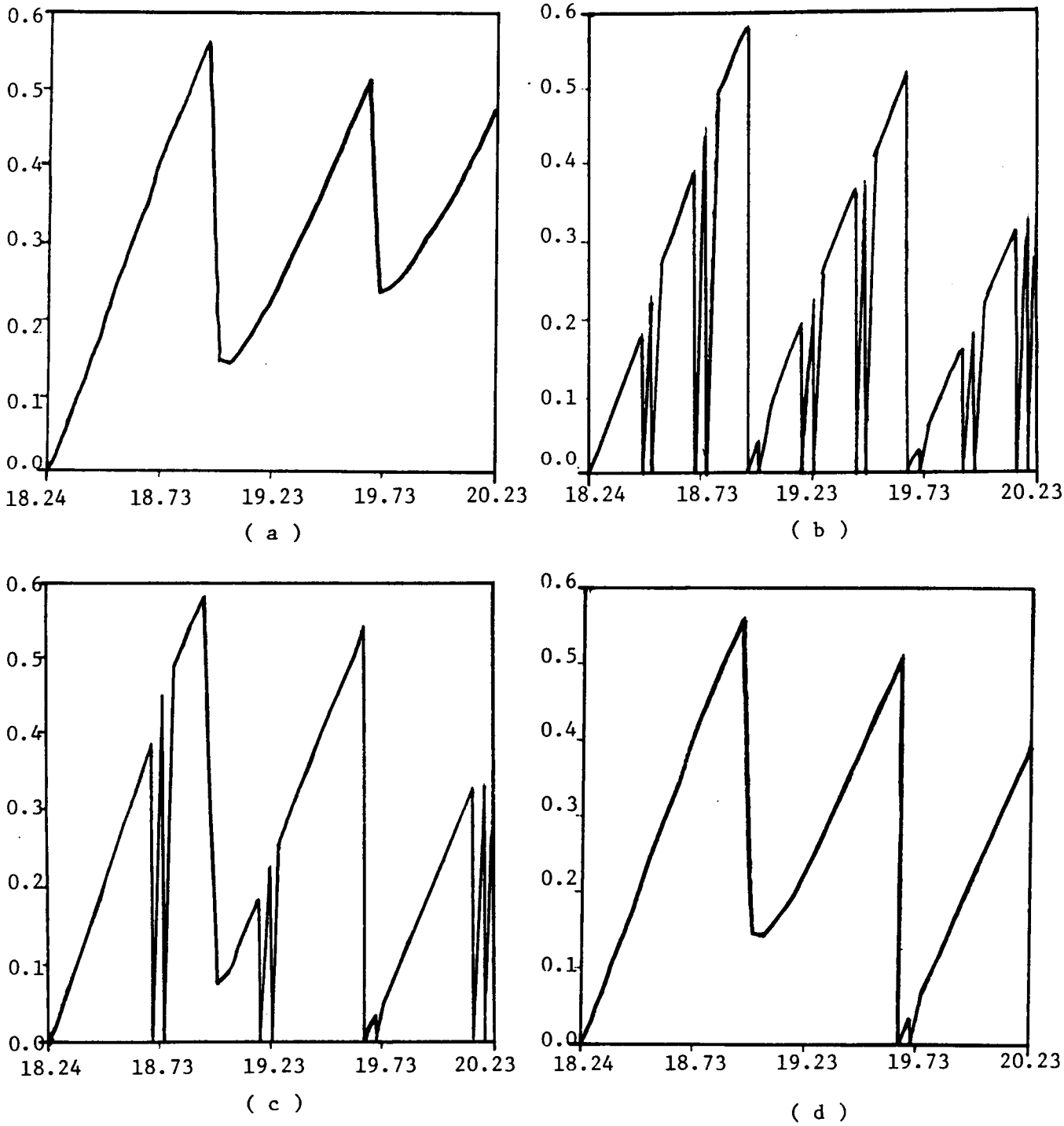


Figure 2.8 NRMS streamfunction errors for the experiments that update the boundary every 15 time steps, but update the interior every
(a) ---------- (b) 5 time steps
(c) 10 time steps (d) 30 time steps

CHAPTER III
THE STATISTICAL APPROACH

III.1 General Objective Analysis

In this section, we derive the general objective analysis formula. This formula is well known and has been used by meteorologists for a long time (Gandin (10)). A recursive algorithm for the general objective analysis is introduced next. This was first discovered by Petersen (20). However, it is not implemented here because not all data are used to interpolate to a grid point in our statistical model. It is included here only for theoretical interest.

Let ψ be a homogeneous random field with zero mean and known covariance function

$$\overline{\psi(x)} = 0 \tag{3.1}$$

$$\overline{\psi(x)\psi(x+\xi)} = C(\xi) \tag{3.2}$$

where overbar denotes ensemble average.

Let

$$\phi(x_i) = \psi(x_i) + e(x_i) \tag{3.3}$$

$$(\ i = 1, \cdots N)$$

be a set of measurements of ψ at the generalized coordinates $X = X_i$ (i = 1, ----N) respectively, Assume the measurement

errors $\epsilon(x_i)$ $(i = 1, \cdots N)$ are uncorrelated with one another and with the field ψ but have known variance E

$$\overline{\epsilon(x_i)\,\psi(x)} = 0 \tag{3.4}$$

$$\overline{\epsilon(x_i)\,\epsilon(x_j)} = E\,\delta_{ij} \quad (i, j = 1, \cdots N) \tag{3.5}$$

We are interested in interpolating $\{\phi(x_i) \mid i = 1, \cdots N\}$ to all points in space in order to reconstruct the whole field. The best linear estimate at a generalized coordinate X

$$\psi^e(x) = \sum_{i=1}^{N} \alpha_i(x)\,\phi(x_i) \tag{3.6}$$

that minimizes the mean square error $E^2(x) = \overline{[\psi^e(x) - \psi(x)]^2}$ has the weighting coefficients

$$\alpha_i^*(x) = \sum_{j=1}^{N} D_{ij}^{-1}\,C(x - x_j) \quad (i = 1, \cdots N) \tag{3.7}$$

where D is a matrix with the (i, j)-th element

$$D_{ij} = C(x_i - x_j) + E\,\delta_{ij} \quad (i, j = 1, \cdots N) \tag{3.8}$$

and the minimal mean square error is

$$E^{*2}(x) = C(0) - \sum_{i,j=1}^{N} C(x - x_i)\,D_{ij}^{-1}\,C(x - x_j) \tag{3.9}$$

This is the famous Gauss-Markov theorem. This powerful technique has been widely used to reconstruct a random field from a set of coarsely distributed data. It is known by many names: objective analysis (Gandin(10)), linear interpolation, extrapolation, and prediction (Petersen(19)), and random field reconstruction (Petersen(18)).

We note in (3.7) that if E ≠ 0, then $\alpha_i^*(x_j) \neq \delta_{ij}$, $E^{*2}(x_i) \neq 0$, $(i, j = 1, \cdots N)$, i.e., the linear least mean square estimate at a data point is not the data at this data point; and the minimal mean square interpolation error at a data point is not identically zero. This is expected since if the data are not perfect, then the estimate at a given data point will be improved by the data at other points. If E = 0, (3.7) and (3.9) become

$$\alpha_i^*(x) = \sum_{j=1}^{N} C_{ij}^{-1} C(x - x_j) \quad (i = 1, \cdots N) \tag{3.7a}$$

$$E^{*2}(x) = C(0) - \sum_{i,j=1}^{N} C(x - x_i) C_{ij}^{-1} C(x - x_j) \tag{3.9a}$$

where C is a matrix with the (i,j)-th element

$$C_{ij} = C(x_i - x_j) \quad (i, j = 1, \cdots N) \tag{3.8a}$$

In this case, we have the following interesting results

$$\alpha_i^*(x_j) = \delta_{ij} \quad (i, j = 1, \cdots N) \tag{3.10}$$

$$E^{*2}(x_i) = 0 \quad (i = 1, \cdots N) \tag{3.11}$$

Namely, the optimal weighting functions $\{\alpha_i^*(x) \mid i = 1, \cdots N\}$ are cardinal and the minimal interpolation error vanishes at data points.

In computing the optimal weighting coefficients and minimal interpolation errors, we have to invert an NXN matrix if there are N data available. That is fine if these data are all that we can get; since we need to invert the NXN matrix only once.

But what happens if one more additional datum arrives? Do we have to recompute the whole thing all over again from the beginning? Set up the (N+1)x(N+1) matrix, invert it, compute the optimal weighting coefficients, the minimal interpolation errors and so on. It would be nice if we could build up on what we have done already, reducing the total computational effort. Fortunately, this turns out to be possible. The following paragraphs describe a recursive algorithm to do the job.

Assume N data $\{\phi(x_i) \mid i = 1, \cdots N\}$ have already been processed; namely, we have computed the optimal weighting functions $\{\alpha_i^N(x) \mid i = 1, \cdots N\}$ and the minimal interpolation error $E_N^2(x)$ for them.

$$\alpha_i^N(x) = \sum_{j=1}^{N} D_{ij}^{-1} C(x-x_j) \quad (i = 1, \cdots N) \tag{3.7b}$$

$$E_N^2(x) = C(0) - \sum_{i,j=1}^{N} C(x-x_i) D_{ij}^{-1} C(x-x_j) \tag{3.9b}$$

Write (3.7b) and (3.9b) in different forms

$$\sum_{j=1}^{N} D_{ij} \alpha_j^N(x) = C(x-x_i) \quad (i = 1, \cdots N) \tag{3.12}$$

$$E_N^2(x) = C(0) - \sum_{i=1}^{N} C(x-x_i) \alpha_i^N(x) \tag{3.13}$$

If one more additional datum $\phi(x_{N+1})$ arrives, the new optimal weighting functions $\{\alpha_i^{N+1}(x) \mid i = 1, \cdots N+1\}$ and the minimal interpolation error $E_{N+1}^2(x)$ satisfy

$$\sum_{j=1}^{N+1} D_{ij} \alpha_j^{N+1}(x) = C(x-x_i) \quad (i = 1, \cdots N+1) \tag{3.14}$$

$$E_{N+1}^2(x) = C(0) - \sum_{i=1}^{N+1} C(x-x_i) \alpha_i^{N+1}(x) \tag{3.15}$$

Consider the matrix

$$\begin{bmatrix} D_{11} & D_{12} & \cdots & D_{1N} & D_{1(N+1)} \\ D_{21} & D_{22} & \cdots & D_{2N} & D_{2(N+1)} \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ D_{N1} & D_{N2} & \cdots & D_{NN} & D_{N(N+1)} \\ D_{(N+1)1} & D_{(N+1)2} & \cdots & D_{(N+1)N} & D_{(N+1)(N+1)} \end{bmatrix} \tag{3.16}$$

In the last row $[\, D_{(N+1)1} \;\; D_{(N+1)2} \cdots D_{(N+1)N} \;\; D_{(N+1)(N+1)}] =$ $[\, C(x_{N+1}-x_1) \;\; C(x_{N+1}-x_2) \cdots C(x_{N+1}-x_N) \;\; C(0)+E]$ and is related to $\{ \alpha_i^N(x_{N+1}) \mid i = 1, \cdots N \}$ through (3.12). $\therefore$ The matrix in (3.16) becomes

$$\begin{bmatrix} D_{11} & D_{12} & \cdots & D_{1N} & C(x_{N+1}-x_1) \\ D_{21} & D_{22} & \cdots & D_{2N} & C(x_{N+1}-x_2) \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ D_{N1} & D_{N2} & \cdots & D_{NN} & C(x_{N+1}-x_N) \\ \sum_{i=1}^{N} D_{i1}\alpha_i^N(x_{N+1}) & \sum_{i=1}^{N} D_{i2}\alpha_i^N(x_{N+1}) & \cdots & \sum_{i=1}^{N} D_{iN}\alpha_i^N(x_{N+1}) & C(0)+E \end{bmatrix} \tag{3.16a}$$

Multiply the i-th row by $-\alpha_i^N(x_{N+1})$ $(i = 1, \cdots N)$ and add them to the (N+1)-th row, we get

$$\begin{bmatrix} D_{11} & D_{12} & \cdots & D_{1N} & C(x_{N+1}-x_1) \\ D_{21} & D_{22} & \cdots & D_{2N} & C(x_{N+1}-x_2) \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ D_{N1} & D_{N2} & \cdots & D_{NN} & C(x_{N+1}-x_N) \\ 0 & 0 & \cdots & 0 & C(0)+E-\sum_{i=1}^{N} C(x_{N+1}-x_i)\alpha_i^N(x_{N+1}) \end{bmatrix} \tag{3.16b}$$

The determinant Δ_{N+1} of (3.16) is equal to the determinant of (3.16b) and is equal to

$$\Delta_{N+1} = \left[C(0) + E - \sum_{i=1}^{N} C(x_{N+1} - x_i) \alpha_i^{N}(x_{N+1}) \right] \Delta_N \tag{3.17}$$

The weighting function $\alpha_{N+1}^{N+1}(x)$ can be calculated from (3.14) by the Cramer's rule

$$\alpha_{N+1}^{N+1}(x) = \frac{1}{\Delta_{N+1}} \begin{vmatrix} D_{11} & D_{12} & \cdots & D_{1N} & C(x-x_1) \\ D_{21} & D_{22} & \cdots & D_{2N} & C(x-x_2) \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ D_{N1} & D_{N2} & \cdots & D_{NN} & C(x-x_N) \\ D_{(N+1)1} & D_{(N+1)2} & \cdots & D_{(N+1)N} & C(x-x_{N+1}) \end{vmatrix} . \tag{3.18}$$

By applying the same strategy to the last row in (3.18), we get

$$\alpha_{N+1}^{N+1}(x) = \frac{\Delta_N}{\Delta_{N+1}} \left[C(x - x_{N+1}) - \sum_{i=1}^{N} C(x - x_i) \alpha_i^{N}(x_{N+1}) \right] \tag{3.18a}$$

From (3.14) and (3.12)

$$\begin{aligned} \sum_{j=1}^{N+1} D_{ij} \alpha_j^{N+1}(x) &= \sum_{j=1}^{N} D_{ij} \alpha_j^{N+1}(x) + D_{i(N+1)} \alpha_{N+1}^{N+1}(x) \\ &= \sum_{j=1}^{N} D_{ij} \alpha_j^{N+1}(x) + C(x_{N+1} - x_i) \alpha_{N+1}^{N+1}(x) \\ &= \sum_{j=1}^{N} D_{ij} \left[\alpha_j^{N+1}(x) + \alpha_{N+1}^{N+1}(x) \alpha_j^{N}(x_{N+1}) \right] = C(x - x_i) \\ &= \sum_{j=1}^{N} D_{ij} \alpha_j^{N}(x) \qquad (i = 1, \cdots N) \end{aligned} \tag{3.19}$$

$$\therefore \quad \alpha_i^{N+1}(x) = \alpha_i^{N}(x) - \alpha_{N+1}^{N+1}(x) \alpha_i^{N}(x_{N+1}) \quad (i = 1, \cdots N) \tag{3.20}$$

From (3.13) and (3.15)

$$E_N^2(x) - E_{N+1}^2(x) = \sum_{i=1}^{N+1} C(x-x_i)\alpha_i^{N+1}(x) - \sum_{i=1}^{N} C(x-x_i)\alpha_i^N(x)$$

$$= C(x-x_{N+1})\alpha_{N+1}^{N+1}(x) + \sum_{i=1}^{N} C(x-x_i)\left(\alpha_i^{N+1}(x) - \alpha_i^N(x)\right)$$

$$= C(x-x_{N+1})\alpha_{N+1}^{N+1}(x) - \sum_{i=1}^{N} C(x-x_i)\alpha_i^N(x_{N+1})\alpha_{N+1}^{N+1}(x)$$

$$= \alpha_{N+1}^{N+1}(x)\left[C(x-x_{N+1}) - \sum_{i=1}^{N} C(x-x_i)\alpha_i^N(x_{N+1})\right]$$

$$= \frac{\Delta_{N+1}}{\Delta_N}\left[\alpha_{N+1}^{N+1}(x)\right]^2 \qquad (3.21)$$

Thus, the recursive procedures for computing $\left\{\alpha_i^{N+1}(x) \mid i=1,\cdots N+1\right\}$ and $E_{N+1}^2(x)$ from $\left\{\alpha_i^N(x) \mid i=1,\cdots N\right\}$ and $E_N^2(x)$ are in order

(i) compute $\left\{\alpha_i^N(x_{N+1}) \mid i=1,\cdots N\right\}$

(ii) compute $\frac{\Delta_{N+1}}{\Delta_N}$ from (3.17)

(iii) compute $\alpha_{N+1}^{N+1}(x)$ from (3.18a)

(iv) compute $\left\{\alpha_i^{N+1}(x) \mid i=1,\cdots N\right\}$ from (3.20)

(v) compute $E_{N+1}^2(x)$ from (3.21)

The logic diagram of the recursive algorithm* which is appropriate for sequential estimation is shown in Figure 3.1. A few interesting remarks follow:

Let

$$\psi_N^e(x) = \sum_{i=1}^{N} \alpha_i^N(x)\,\phi(x_i)$$

$$\psi_{N+1}^e(x) = \sum_{i=1}^{N+1} \alpha_i^{N+1}(x)\,\phi(x_i)$$

The difference between the optimal linear estimates from (N+1) data and N data at a general coordinate x is

---

* This recursive algorithm is taken from Petersen (20). I supplied the details of derivation here. In practice, we work only in the low ranges of the parameter N (see Section III.3), so it is not very useful for our purpose. Also, this algorithm is not a very good algorithm as pointed out to me by Prof. Anderson.

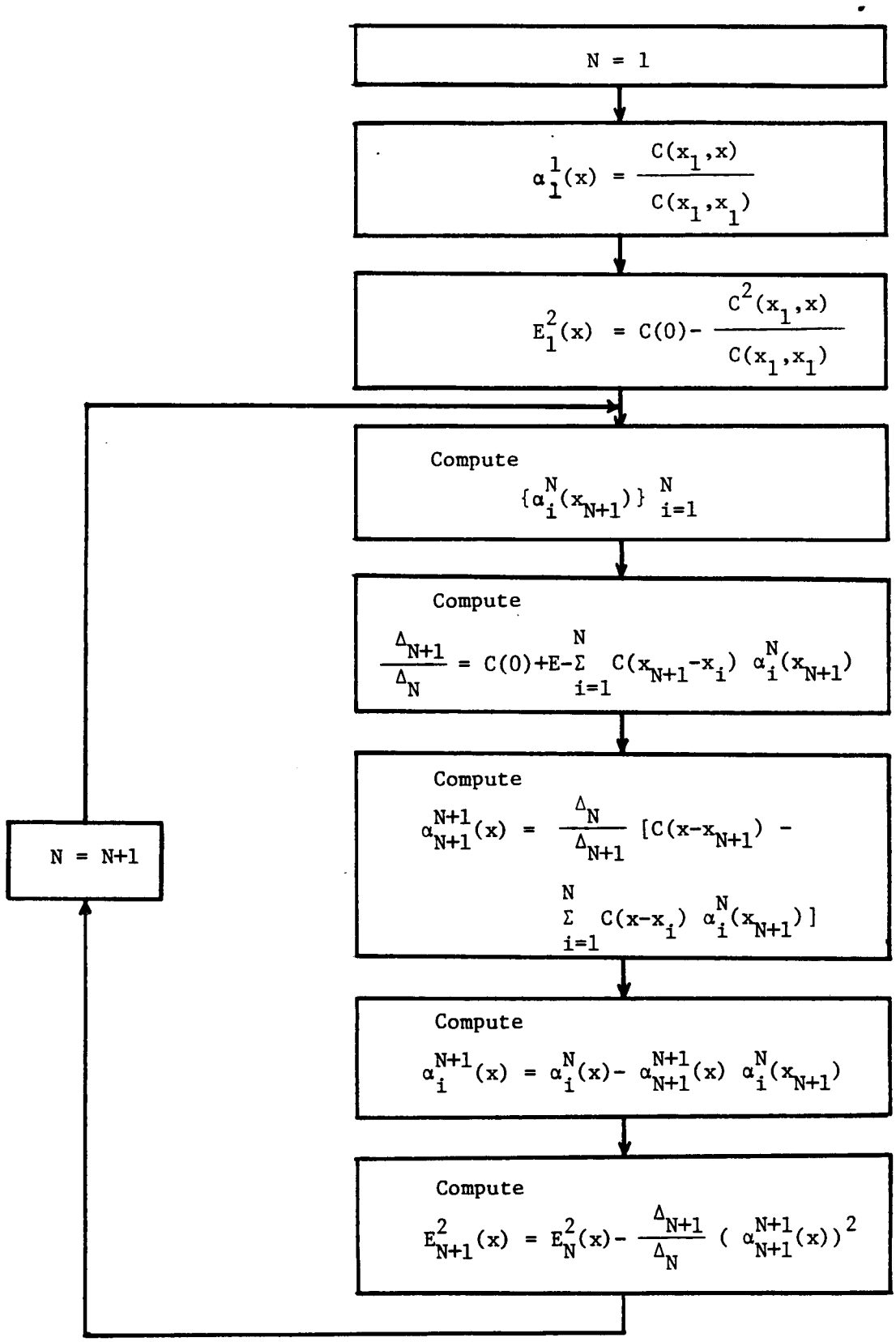


Figure 3.1 A recursive algorithm for the general objective analysis.

$$\psi^e_{N+1}(x) - \psi^e_N(x) = \sum_{i=1}^{N+1} \alpha_i^{N+1}(x)\,\phi(x_i) - \sum_{i=1}^{N} \alpha_i^N(x)\,\phi(x_i)$$

$$= \alpha_{N+1}^{N+1}(x)\,\phi(x_{N+1}) + \sum_{i=1}^{N} \left[\alpha_i^{N+1}(x) - \alpha_i^N(x)\right]\phi(x_i)$$

$$\overset{(3.20)}{=} \alpha_{N+1}^{N+1}(x)\,\phi(x_{N+1}) - \sum_{i=1}^{N} \left[\alpha_{N+1}^{N+1}(x)\,\alpha_i^N(x_{N+1})\right]\phi(x_i)$$

$$= \alpha_{N+1}^{N+1}(x)\left[\phi(x_{N+1}) - \sum_{i=1}^{N} \alpha_i^N(x_{N+1})\,\phi(x_i)\right]$$

$$= \alpha_{N+1}^{N+1}(x)\left[\phi(x_{N+1}) - \psi^e_N(x_{N+1})\right]$$

(3.22)

So it is equal to the difference between the actual and the estimated value of the (n+1)th observation, multiplied by the weighting coefficient at X from the (n+1)th observation.

From (3.21)

$$E_N^2(x) - E_{N+1}^2(x) = \frac{\Delta_{N+1}}{\Delta_N}\left[\alpha_{N+1}^{N+1}(x)\right]^2 \qquad (3.21)$$

And the minimal interpolation error always decreases as new data are added. $\frac{\Delta_{N+1}\left[\alpha_{N+1}^{N+1}(x)\right]^2}{\Delta_N}$ can be regarded as a measure of the "figure of merit" of the new data point at a generalized coordinate X.

III.2 The Statistical Model

If we let the generalized coordinate X in (3.6) and (3.7) be (r,t), where r is the spatial coordinate, t is the time, then they become the space-time objective analysis formulas. These formulas can be used to interpolate or extrapolate in space a patch of data to points within the patch or outside the patch at a given time, or else can be used to predict the

field at a future time. The space-time interpolation scheme which uses the data at the same time level and forecasts the field forward in time forms our statistical forecast model.

The quality of our interpolation and forecast results depends strongly on how well the modelling assumptions are satisfied by the available data. If these assumptions are not satisfied, some results we have obtained before will not be valid. For example, (3.21) states that the minimal mean square interpolation error will decrease as the number of data available increases. This result is not always true in real situations. The data added may cause the result to deteriorate instead of improving it. This occurs sometimes in our simulation experiments.

One of the potentially violable assumptions we have made is that the correlation function is homogeneous, stationary , and does not depend on the strength of the field. These are very restrictive assumptions. Very few sets of real data satisfy these assumptions exactly. Of course we could have made more general assumptions which allow nonhomogeneity and nonstationarity. However, if the random field is not homogeneous and stationary, then the correlation function has to be known at every point and at every time. This is a great deal of information to compute or estimate if an a priori analytic correlation function is not available. The assumptions we made here were reached after taking the computational and accuracy requirements into consideration.

The correlation function which we employ in the statistical model is computed from the simulated ocean data set provided by the exterior calculation of the dynamical model. Since the correlation is assumed to be homogeneous and stationary, it is computed via the formula

$$C(\Delta x, \Delta y, \Delta t) = \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=t_0}^{t_f} \psi(x_i+\Delta x, y_i+\Delta y, t+\Delta t)\,\psi(x_i, y_i, t) \quad (3.22)$$

where t_o = period two, t_f = period six - Δ t, since only the data from period two to period six are stored from the exterior calculation. (3.22) is the average of all products of two data samples which are separated by $(\Delta x, \Delta y)$ in space and have a time lag Δ t. This correlation function has been computed for all

$$-125\,km \le \Delta x, \Delta y \le 125\,km \quad (3.23)$$

in increment of a grid interval (15.625 km)

$$0 \le \Delta t \le 48 \text{ time steps} \quad (3.24)$$

in increment of a time step (1.008 day), and are stored for later use. Some of the correlation function maps computed are shown in Figure 3.2. Figure 3.1a-f are the correlation function maps for Δt = 0, 8, 16, 24, 32, 40 time steps respectively. The size of the domain shown is 250 km square.

Although we have computed the correlation function for $\Delta t \ge 0$, the correlation function for Δ t < 0 can be inferred from this due to the assumptions of homogeneity and stationarity

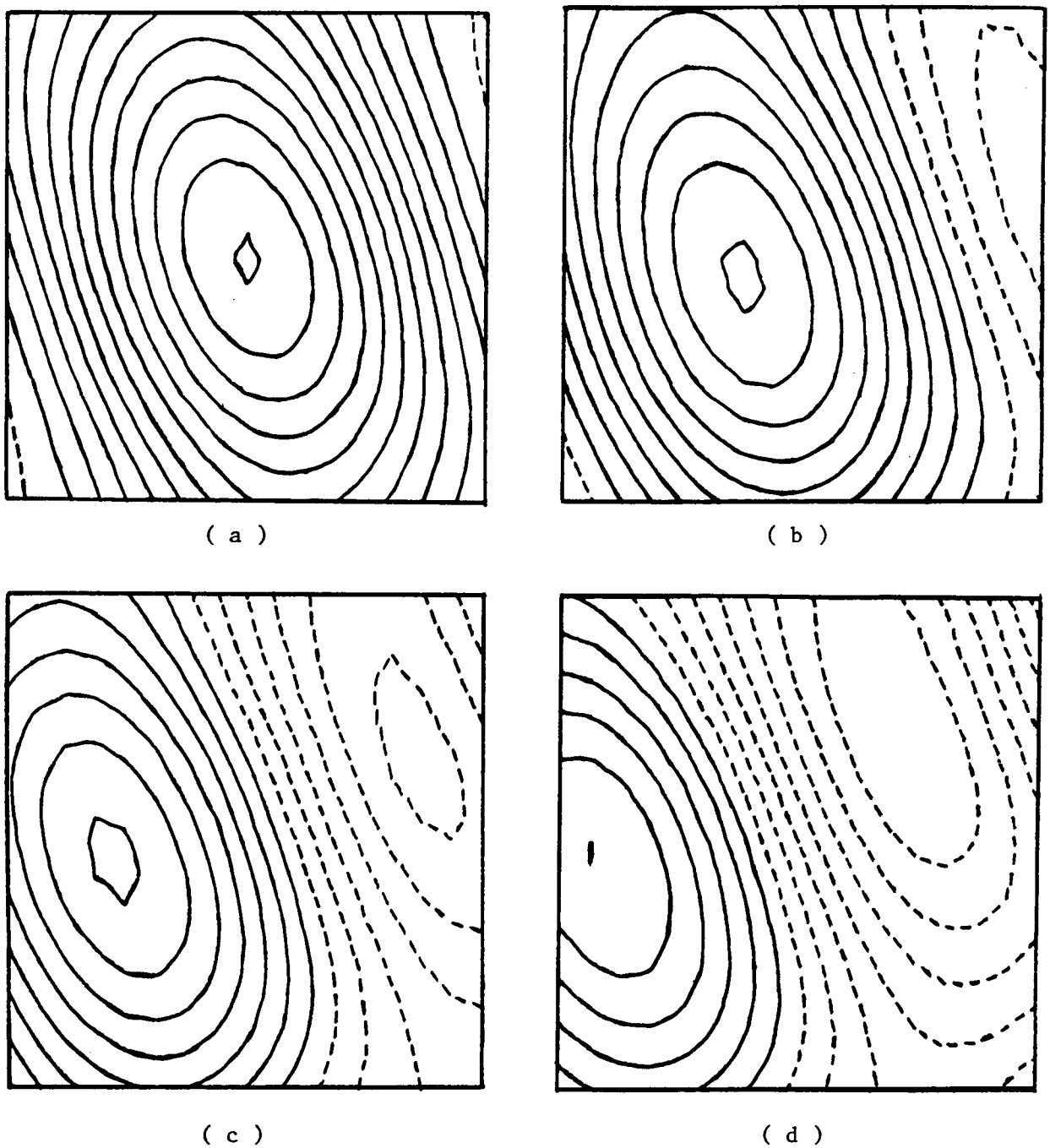
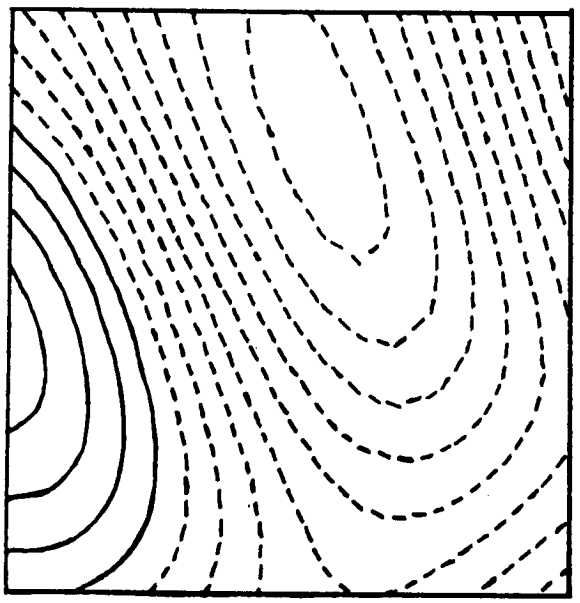
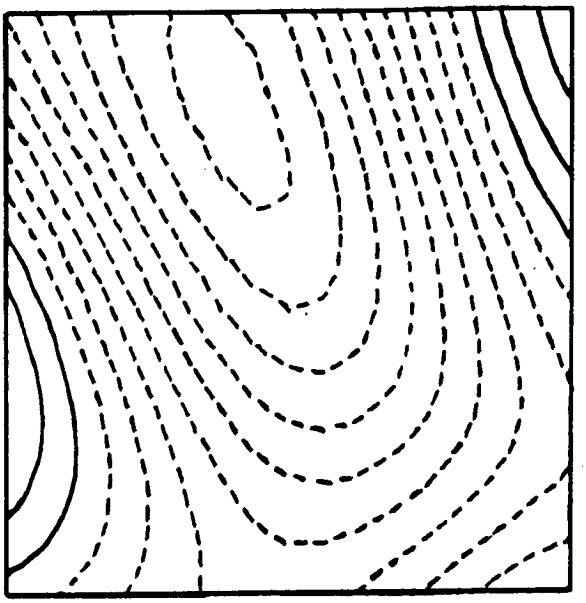


Figure 3.2



(e)



(f)

Figure 3.2 (Cont.)
The calculated correlation function maps for Δt =
(a) 0 (b) 8
(c) 16 (d) 24
(e) 32 (f) 48
time steps.
The maps are shown for -125 km $\leq \Delta x, \Delta y \leq$ 125 km.

$$C(\Delta x, \Delta y, \Delta t) = \overline{(\psi(x+\Delta x, y+\Delta y, t+\Delta t)\,\psi(x,y,t))}$$

$$= \overline{(\psi(x+\Delta x, y+\Delta y, t+\Delta t)\,\psi(x+\Delta x-\Delta x, y+\Delta y-\Delta y, t+\Delta t-\Delta t))}$$

$$= \overline{(\psi(x,y,t)\,\psi(x-\Delta x, y-\Delta y, t-\Delta t))}$$

$$= C(-\Delta x, -\Delta y, -\Delta t)$$

(3.25)

It can be seen that the correlation function reveals a westward propagating property which is typical of mid-latitude mid-ocean eddies.

We now discuss the statistical forecast scheme in detail. In interpolating from a collection of data to a given point by the space-time objective analysis, we found that it is very inefficient if all data are going to be used. Too much time has to be spent to invert a large matrix and to compute the weighting coefficients in (3.7). Thus, we limit the number of data points chosen to N (a parameter). Only those points which have the first N highest correlations with the given point are chosen. This can increase the efficiency very drastically if N is much smaller than the original number of data. Another difficulty arises when the data points chosen are too highly correlated with each other. In this case, the covariance matrix becomes ill-conditioned and we incur a serious numerical diffi-

culty. To avoid this problem, the data points are chosen in a way such that no two data chosen have a correlation higher than C_{max} (a parameter). This is only an empirical precautionary procedure and is not theoretically justified. Finally, since the correlation functions are only computed for $-125\ km \leq \Delta x, \Delta y \leq 125\ km$, only those data points with distances among them satisfying $-125\ km \leq \Delta x_{ij}, \Delta y_{ij} \leq 125\ km$, will be chosen, where ΔX_{ij} is the distance in the X-direction between data point i and data point j.

Taking all the above considerations into account, our data points selecting algorithm goes like this.

(i) Sort all the data points according to their correlation with the point to be interpolated. (A shell sorting algorithm is used in the statistical model - See Knuth (15)).

(ii) Choose the data point which has the highest correlation.

(iii) Search along the sorted list for a data point such that it is within 125 km square and 48 time steps of all the previous ones chosen and it does not have correlation more than C_{max} with all the previous ones chosen.

(iv) If the number of the chosen data points is equal to N, or all the available data have been run out, then stop, otherwise go to (iii).

This algorithm is represented in Figure 3.3.

After the data points have been chosen according to (i)-(iv). The standard objective analysis routine is invoked to do the interpolation using these chosen data points. This completes

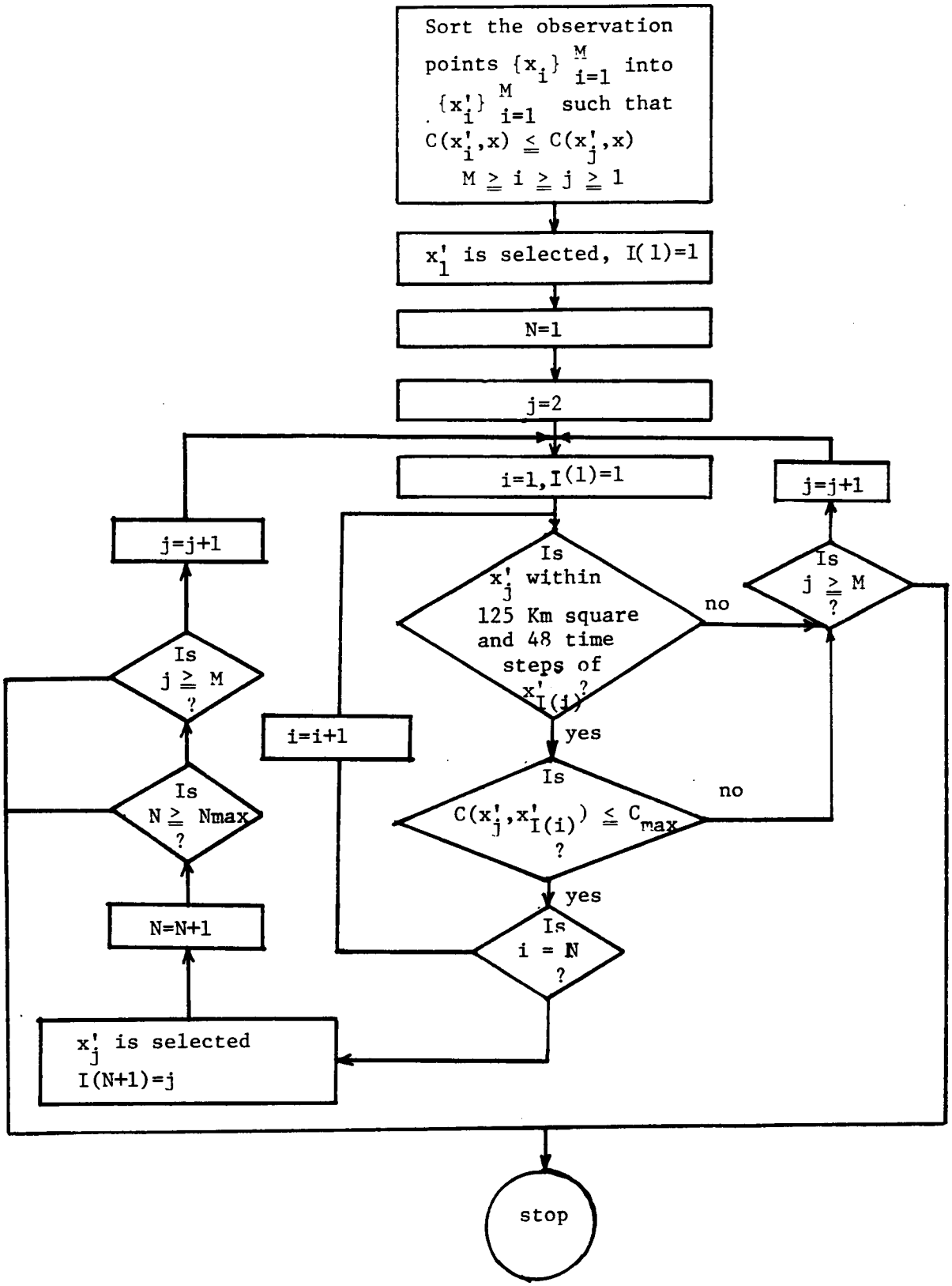


Figure 3.3 The data points selection algorithm in the statistical model.

the detailed description of our statistical model.

III.3 Parameters Study and Statistical Forecast Examples

In our description of the statistical model in Section III.1, we introduced two parameters whose values were to be chosen judiciously to produce accurate, efficient, and stable results. These parameters are N (the number of data points to be chosen), and C_{max} (the maximum correlation any two chosen data points can have). Ideally we could use analytic methods to find the optimal parameters once and for all. However, as we pointed out earlier, the assumptions we have made in our statistical model may not correctly represent the real situation so well. Any result derived theoretically from these assumptions may be misleading. Thus, we resort to empirical studies in choosing the right parameters here. We run several simulation experiments which use various combinations of parameters and see which combination produces the best results. We then try to give some explanation of the results if possible.

We use the statistical forecasts which employ data at the same time level for our parameters study here. Some other parameters are also introduced. Let S and Fil be the parameters that represent the sampling of the initial field and the filter used. S = n means that the initial field is sampled at every n grid points. Fil = (m,n) means that an m-th order Shapiro filter* is applied n times after the field has been predicted. Of course not all of these parameters are independent. For example, the actual number of points chosen depends on how the parameters

* Shapiro (21), (22).

C_{max} and S are chosen. If C_{max} is chosen too small and/or S is chosen too large, then N can't exceed a certain number. All the experiments performed in this regard are summarized in Table 3.1.

Expt. No.	S	N	C_{max}	Filter order	Filter times	NRMS streamfunction errors after (in percentage) 5 (time steps)	10	15
3.1	1	1	---	4	1	18.6	39.4	64.4
3.2	1	5	0.99	-	-	24.0	42.2	66.4
3.3	1	5	0.98	4	1	23.5	41.7	63.6
3.4	1	5	0.98	2	1	23.5	41.7	63.5
3.5	1	5	0.98	2	10	23.6	41.7	63.3
3.6	1	5	0.97	2	1	24.0	43.2	60.4
3.7	1	5	0.97	2	10	24.2	43.2	60.0
3.8	1	5	0.95	2	1	25.0	43.4	65.5
3.9	1	5	0.95	2	10	25.2	43.5	65.6
3.10	1	7	0.91	2	10	24.1	43.0	59.0
3.11	1	10	0.99	-	-	24.3	45.5	60.9
3.12	1	10	0.98	2	1	27.7	85.3	169.0
3.13	1	10	0.98	2	10	25.7	67.0	131.0
3.14	1	10	0.96	2	1	245.3	332.4	145.2
3.15	1	10	0.96	2	10	215.0	299.7	142.3
3.16	1	15	0.98	2	1	150.9	56.9	69.9
3.17	1	15	0.98	2	10	124.0	52.5	62.7
3.18	2	1	---	4	1	24.0	42.4	61.9
3.19	2	2	0.91	4	1	24.3	41.8	61.4
3.20	2	3	0.91	4	1	23.6	41.8	61.9
3.21	2	4	0.91	4	1	23.9	41.9	61.0
3.22	2	5	0.98	4	1	24.2	42.5	60.3
3.23	2	10	0.9	4	1	35.8	50.3	70.5
3.24	3	1	---	4	1	28.3	43.8	64.7
3.25	3	5	0.98	4	1	24.2	42.0	59.6
3.26	4	5	0.98	4	1	25.1	42.3	60.5
3.27	4	5	0.98	4	2	25.1	42.3	60.4
3.28	4	5	0.98	2	1	25.0	42.2	60.2
3.29	4	5	0.98	2	2	24.9	42.1	60.0
3.30	4	5	0.98	2	10	24.7	42.0	59.4

Table 3.1 Summary of the experiments performed in the parameters study

From Table 3.1, we observe the following:

(a) If we have a very dense initial field of data, say S = 1, then we don't need too many data points to forecast the field in

the future. As a matter of fact, a single data point of the initial field, one which is best correlated with the point to be interpolated can be enough. And more data points added will not necessarily improve the result, as (3.21) indicates. This is strong evidence that the model assumptions are not satisfied too well. As a matter of fact, the more data we have, the more difficult the computation will be, as Experiments 3.14-3.17 show. This can be explained roughly as follows. The covariance matrix (3.8) has the form (after dividing each row by $C(0)+E$)

$$\begin{bmatrix} 1 & \alpha & \beta & \cdot & \cdot & \cdot \\ \alpha & 1 & \gamma & \cdot & \cdot & \cdot \\ \beta & \gamma & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 \end{bmatrix} \tag{3.26}$$

If S = 1 and C_{max} is not set too small, then $\alpha, \beta, \gamma \cdots$ are numbers which are very close to 1. For simplicity, we consider the matrix

$$\begin{bmatrix} 1 & \alpha & \alpha & \cdot & \cdot & \cdot \\ \alpha & 1 & \alpha & \cdot & \cdot & \cdot \\ \alpha & \alpha & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 \end{bmatrix} \tag{3.26a}$$

The determinant of the matrix is $(1-\alpha)^{N-1}[1+(N-1)\alpha]$. This determinant is shown in Figure 3.4 as a function of N parameterized by α . We can see as N and/or α increases, the determinant decreases and has a tendency to destabilize the numerical computation. Decreasing N and/or α ($\sim C_{max}$) has a stabilizing

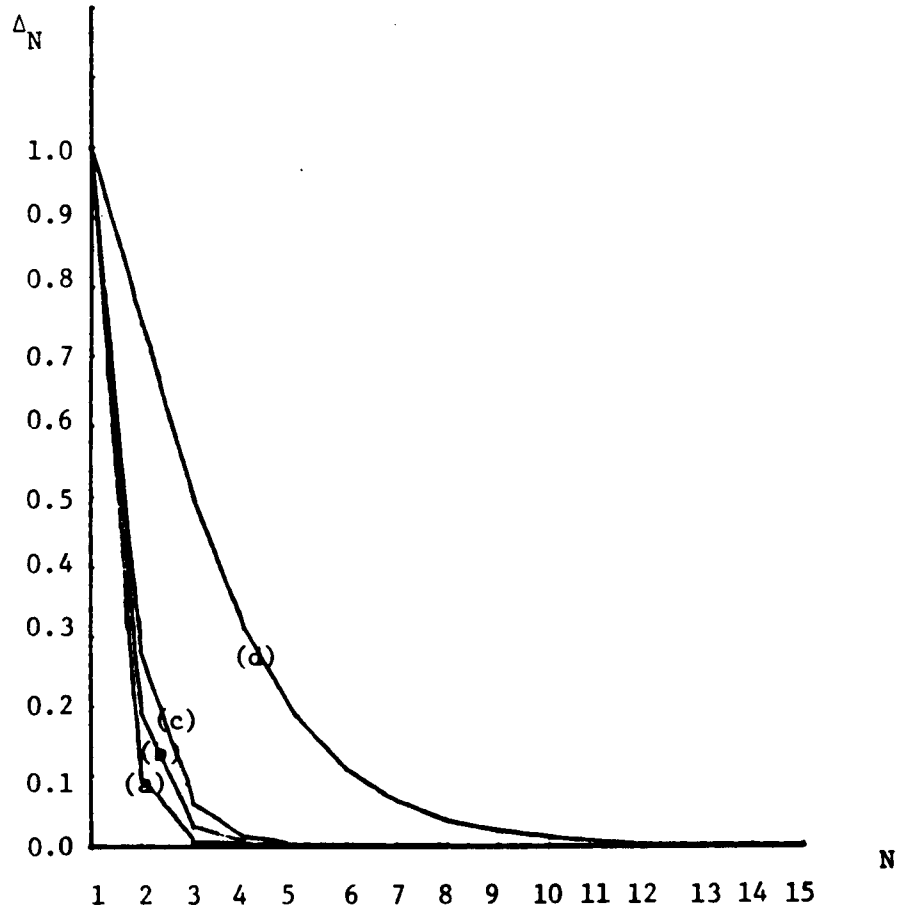


Figure 3.4 The determinant $\Delta_N = (1-\alpha)^{N-1}[1+(N-1)\alpha]$ as a function of N for α =

(a) 0.95 (b) 0.90 (c) 0.85 (d) 0.50

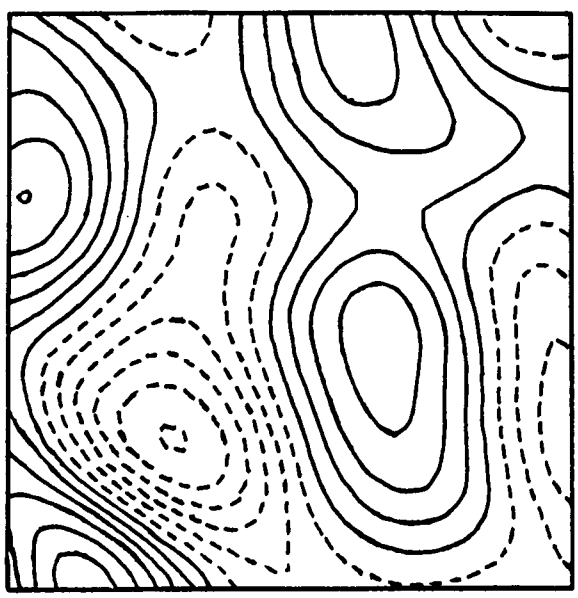
effect.

(b) For higher S, increasing N can improve the result. This is expected, since, if we have a coarse sampling, taking the average of the values of more data points will be better than just using the value of the most correlated data point. For S = 2, N can be chosen to be 3. For S = 3, N = 5. In general, N = f(s). This function has to be determined empirically for every S. But this won't bother us, because only small values of S are most frequently used.

(c) Increasing the order and the number of times of applying the Shapiro filter can improve the result. This is because the forecast field is not very smooth, and any way of smoothing the field can decrease the error.

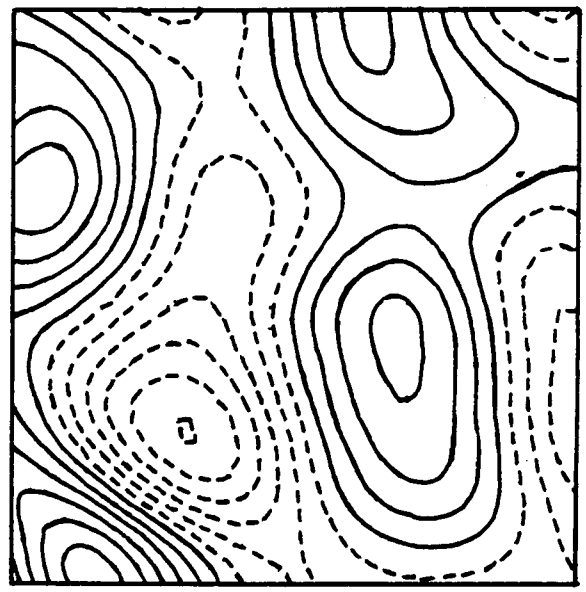
One of the statistical forecast examples is shown on Figure 3.5. Using the true data at period 4.5, we statistically forecast the field forward in time. Figure 3.5a-h shows the result after 2, 4, 6, 8, 10, 12, 14, 16 time steps respectively. The figures on the left hand side are the true fields, while the figures on the right hand side are the statistically forecast fields. The NRMS error curve of the statistical forecast is shown in Figure 3.6. The statistical forecast is found to reproduce the true field very well except near the eastern boundary. This is because the information expected to come in is missing, so the field near the eastern boundary is simply obtained from the old boundary data with decay in space and time.

This can be explained by a one dimensional example. Consider

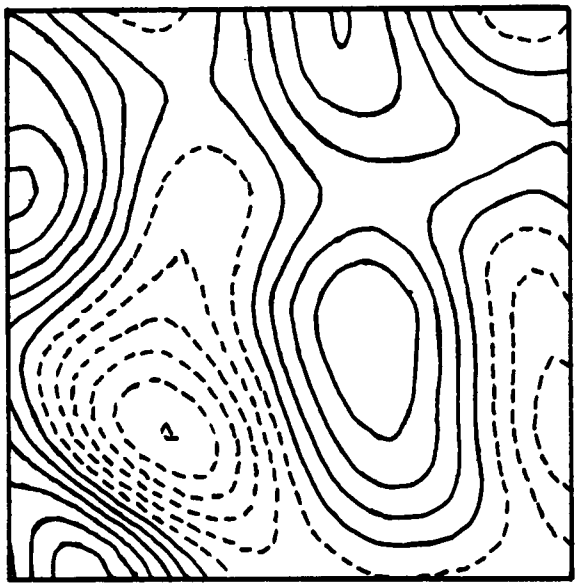


Min=-6.1 Max=5.6 CI=1.0

(a)

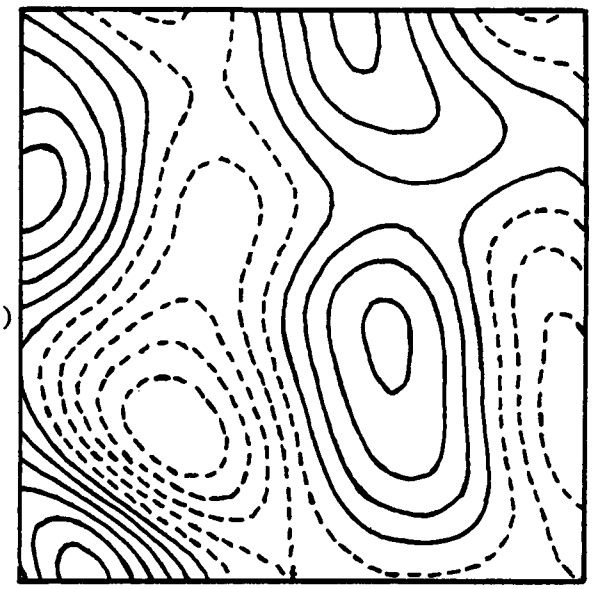


Min=-6.0 Max=5.5 CI=1.0

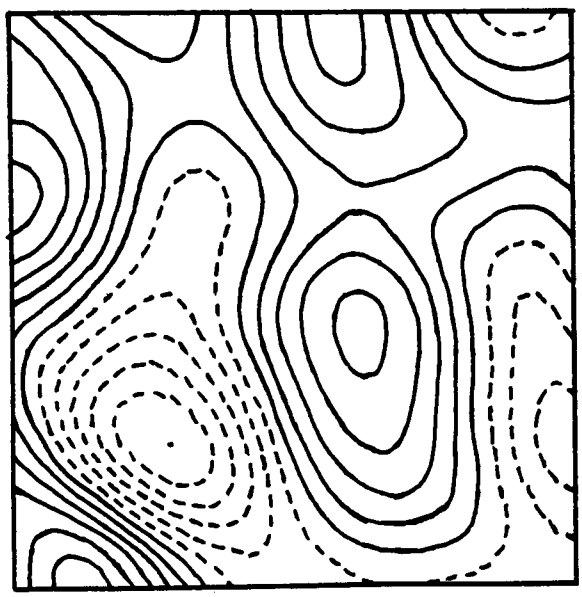


Min=-6.1 Max=5.7 CI=1.0

(b)

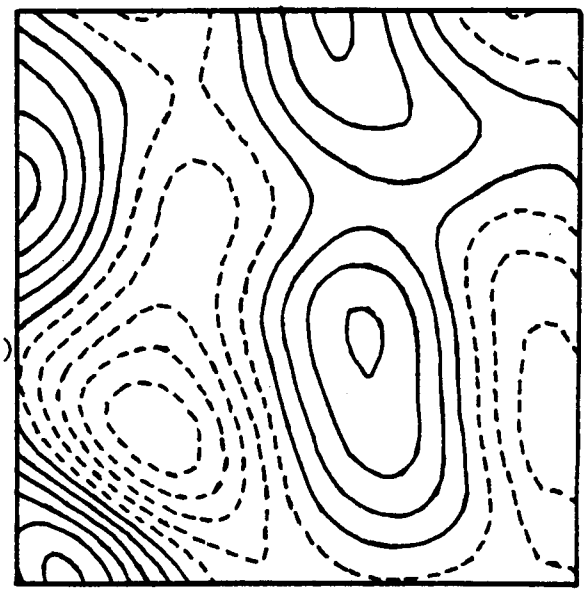


Min=-6.0 Max=5.5 CI=1.0



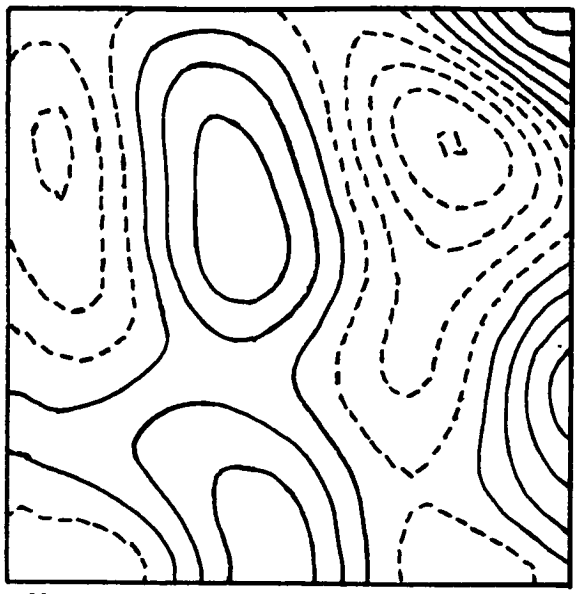
Min=-6.0 Max=5.7 CI=1.0

(c)



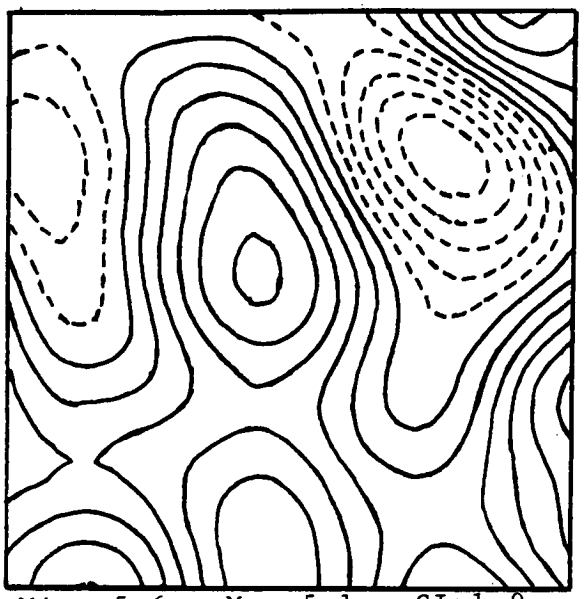
Min=-5.8 Max=5.3 CI=1.0

Figure 3.5

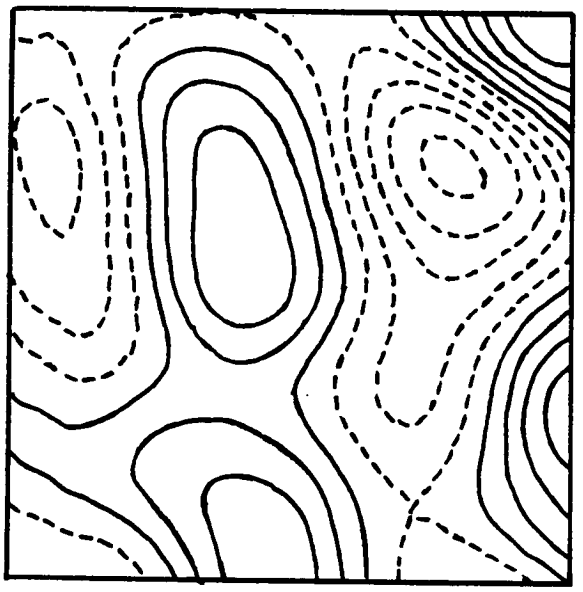


Min=-6.0 Max=5.7 CI=1.0

(d)

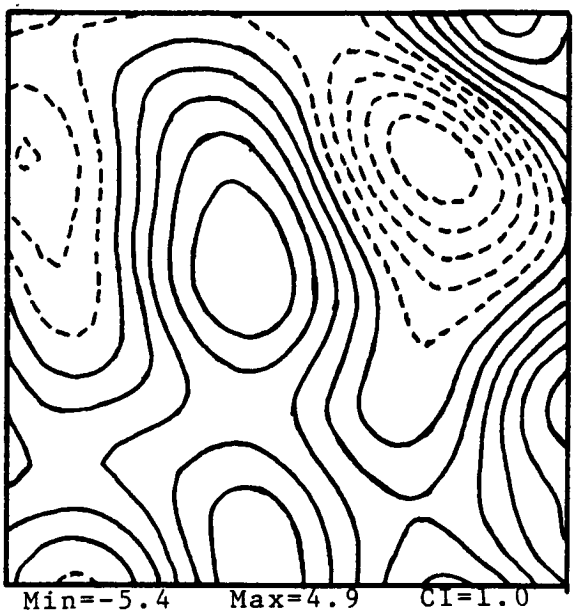


Min=-5.6 Max=5.1 CI=1.0

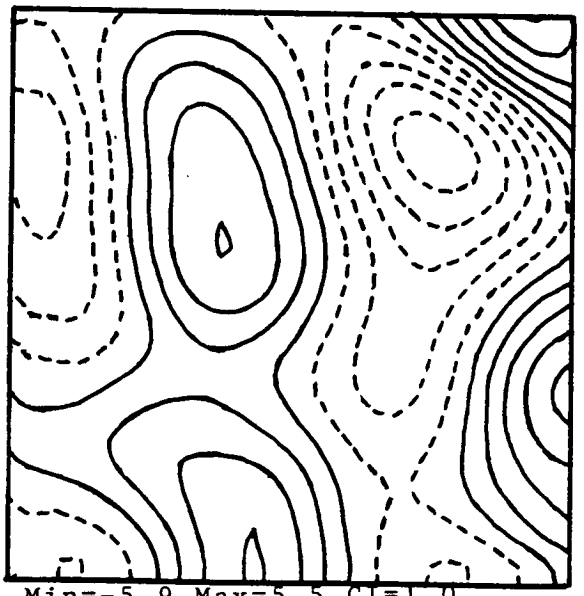


Min=-5.9 Max=5.6 CI=1.0

(e)

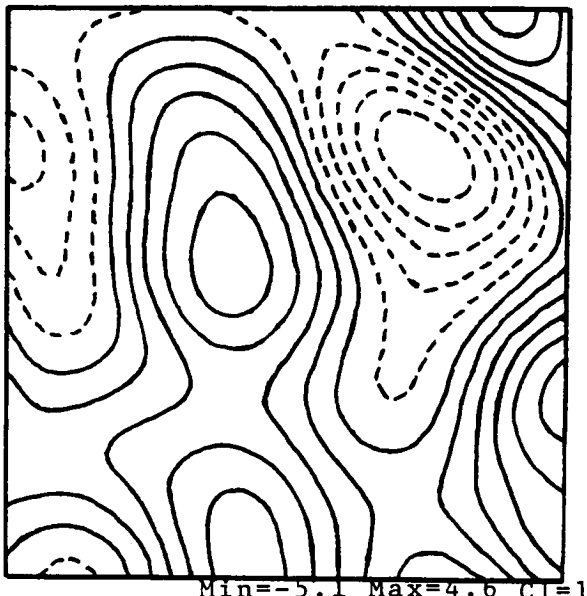


Min=-5.4 Max=4.9 CI=1.0



Min=-5.9 Max=5.5 CI=1.0

(f)



Min=-5.1 Max=4.6 CI=1.0

Figure 3.5 (Cont.)

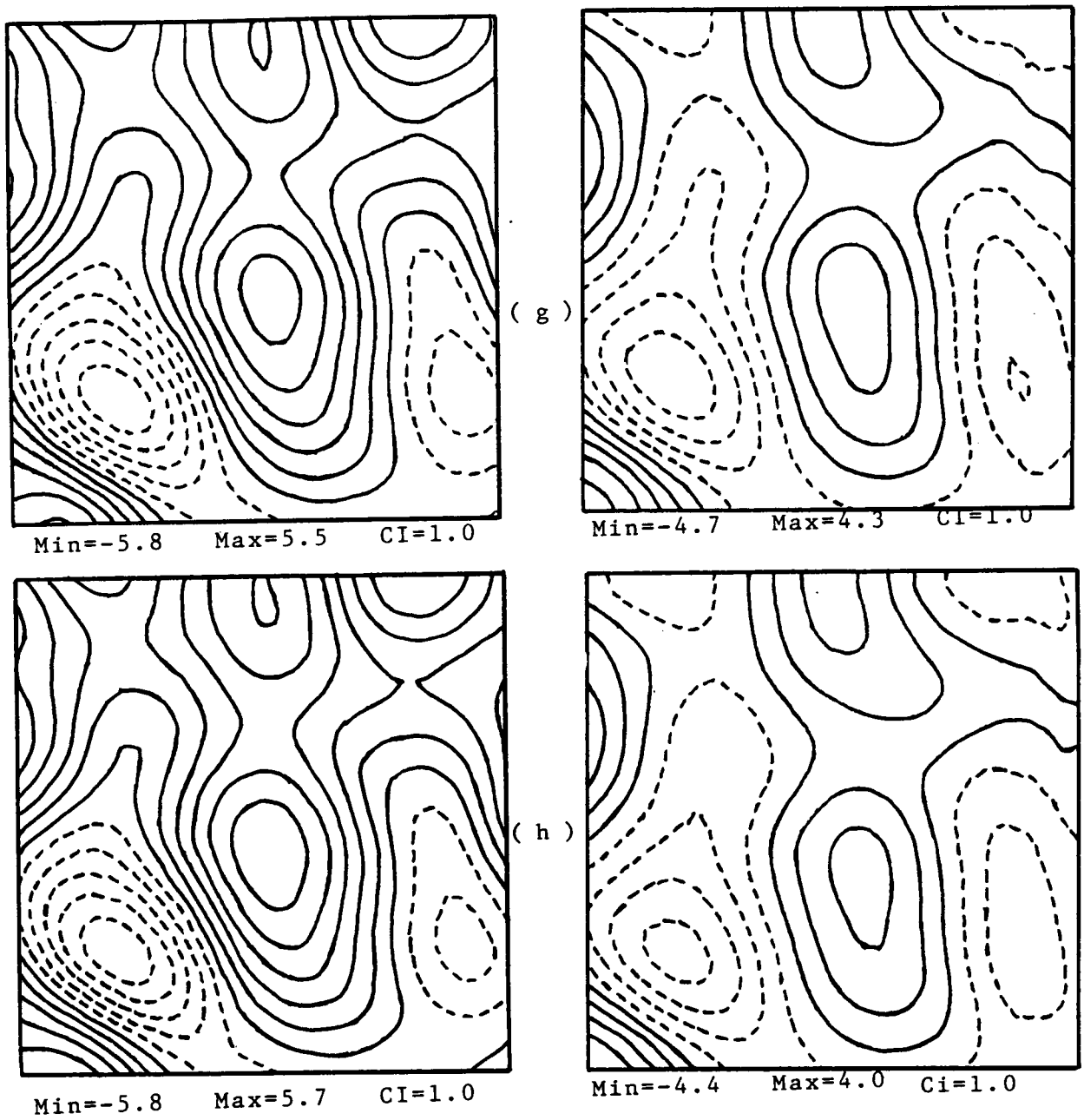


Figure 3.5 (Cont.)
Statistical forecast results using the true field at period 4.5 as the observation data after
(a) 2 (b) 4 (c) 6
(d) 8 (e) 10 (f) 12
(g) 14 (h) 16 time steps.
The figures on the left hand side are the true fields, while the figures on the right hand side are the statistically forecast fields.

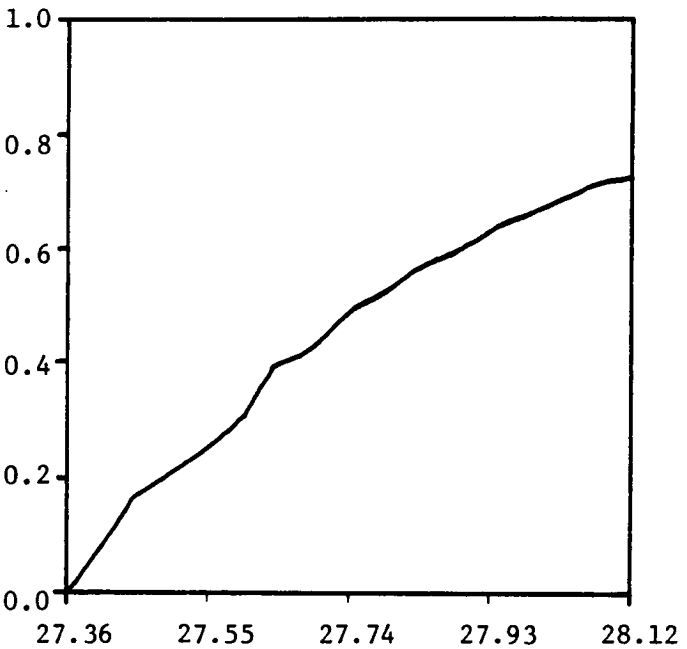


Figure 3.6 The NRMS streamfunction error for the statistical forecast that uses the true data at period 4.5.

a true field

$$\psi^t(x,t) = \cos(x+ct+\theta_0) \quad -L \le x \le L \tag{3.27}$$

in an open domain (-L, L). The correlation function is

$$C(\Delta x, \Delta t) = \frac{1}{2}\cos(\Delta x + c\Delta t) \tag{3.28}$$

If we use only the best correlated point in the space-time interpolation formula to forecast the field $\cos(x+\theta_0)$ at time 0 forward in time, the statistical forecast field will be

$$\psi^s(x,t) = \begin{cases} \cos(x+ct+\theta_0) & -L \le x \le L-ct \\ \cos(x+ct-L)\cos(L+\theta_0) & L-ct \le x \le L \end{cases} \tag{3.29}$$

for $0 \le t \le \frac{\pi}{2c}$. The field between L-ct and L is obtained from the boundary data at time 0 $\cos(L+\theta_0)$ multiplied by a decaying factor $\cos(x+ct-L)$ in space and time.

The difference between the statistical forecast and the true field in this region is

$$\begin{aligned} \psi^e(x,t) &= \psi^s(x,t) - \psi^t(x,t) \\ &= \cos(x+ct-L)\cos(L+\theta_0) - \cos(x+ct+\theta_0) \\ &= \cos(x+ct-L)\cos(L+\theta_0) - \cos(x+ct-L)\cos(L+\theta_0) \\ &\quad + \sin(x+ct-L)\sin(L+\theta_0) \\ &= \sin(x+ct-L)\sin(L+\theta_0) \end{aligned} \tag{3.30}$$

This is illustrated in Figure 3.7.

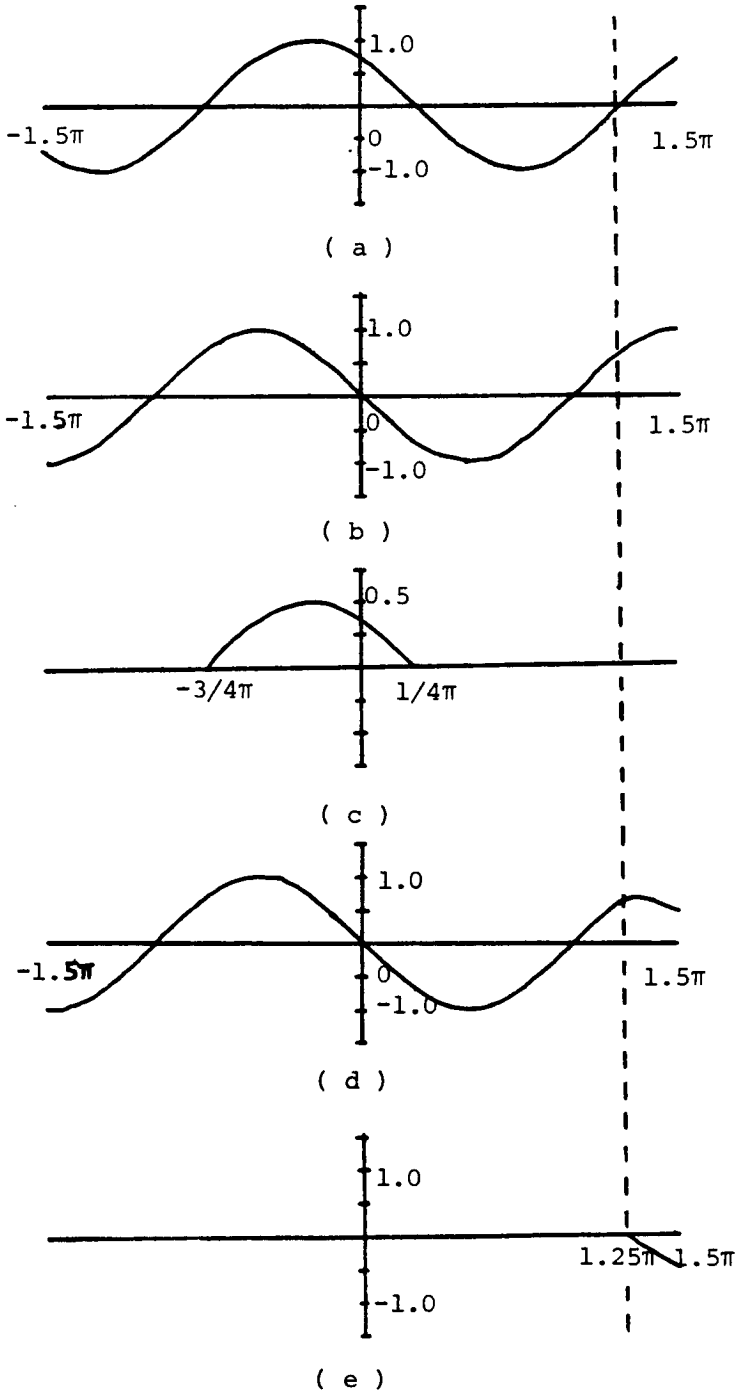


Figure 3.7 A one dimensional example illustrating the space-time objective analysis using the best correlated point.
(a) The true field at time 0: $\cos(x+1/4\pi)$.
(b) The true field at time $ct=1/4\pi$: $\cos(x+1/2\pi)$.
(c) The correlation map for $c\Delta t=1/4\pi$: $1/2\cos(\Delta x+1/4\pi)$.
(d) The statistical forecast field using the space-time objective analysis.
(e) The statistical forecast error field.

III.4 Statistical Forecasts using Data at several Time Levels

In this section, we give some statistical forecast simulation results in which data at several time levels are used. This reflects the situation of a ship steaming in the mid-oceans and collecting the data along its path. We are interested in finding an optimal path of collecting data in order to construct the initial condition for the dynamical model. This is close to the classical problem of finding the most efficient distribution of meteorological stations to collect the Atmospheric data (Gandin [10], Chapter 7), but not quite. The data collecting is constrained by the ship's movement. The speed of the ship is finite, so the data collected are in a sequential order and are not simultaneous.

We can formulate this problem as an optimal control problem. The problem is to find the optimal cruising path and data sampling scheme. The objective is to minimize the error of the constructed field, as well as the total length of the cruising path, and the total number of measurements (such as the number of XBT's cast, etc.), when subjected to the constraints that the data are collected on a feasible cruising path (here feasible means that the speed limit of a ship is not violated). To solve this optimal control problem analytically is very difficult. Although the optimal solution can be solved analytically by making further assumptions about the problems, it may not be practically useful since these assumptions may

not be satisfied realistically. We have this experience in Section III.2.

Here we devise a cruising path based on our intuition. We want to see how the space-time objective analysis performs in this case. Of course this cruising path is far from optimal. However, this is the first step in attempting to solve the problem. Eventually we may have to resort to a simulation method to find the optimal cruising path, just as we have done in Section III.2 to find the optimal parameters.

The one we consider here is to start the cruising and data collecting from the eastern boundary. (This is based on the intuition that the waves propagate westwards. The information that comes in from the eastern boundary stays the longest in the domain.) We steam the ship along the eastern boundary collecting data at every 4 grid points (62.5 km). This is accomplished in one day. The speed of the ship must therefore exceed $\frac{500}{24}$ = 21 km/hr, which is achievable by today's standard. The second day we steer the ship along the path which is 4 grid points west of the eastern boundary, collecting the data every 4 grid points. The process continues on. On day 10, we return the ship to the beginning of the eastern boundary without collecting any data and repeat the cruising and data collecting processes. The cruising and data collecting path is shown in Figure 3.8. The statistical forecast results are shown in Figure 3.9. Figure 3.9a-s represents the true field and forecast field on day 1-19 respectively. The true fields are shown

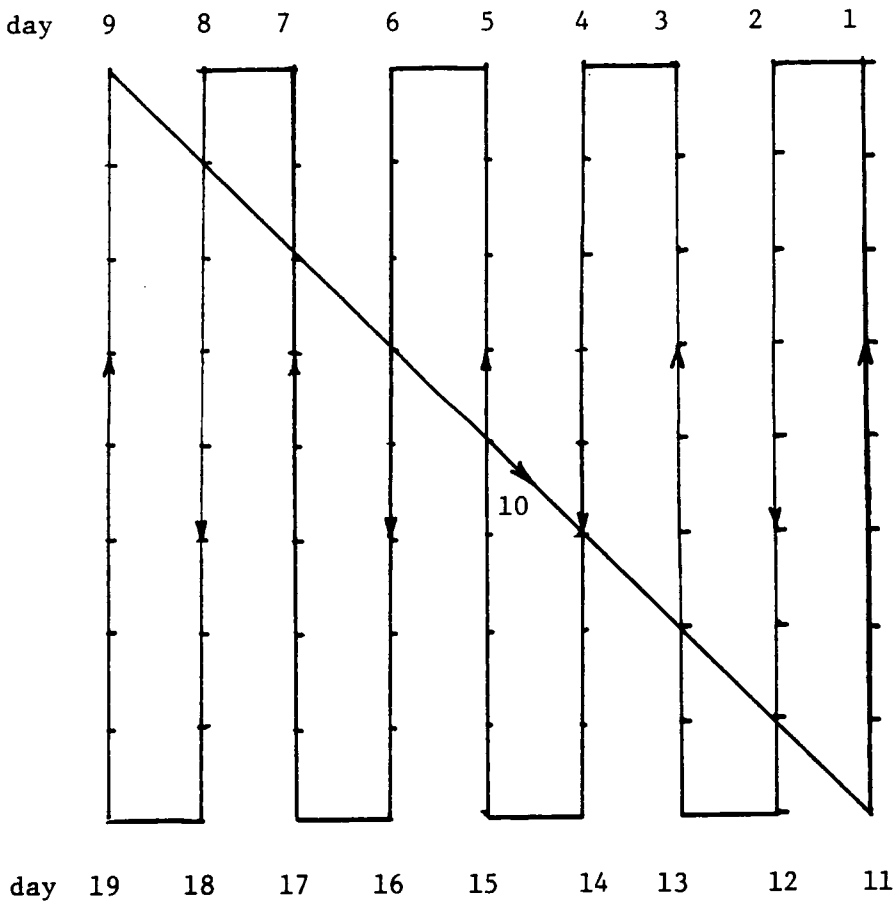
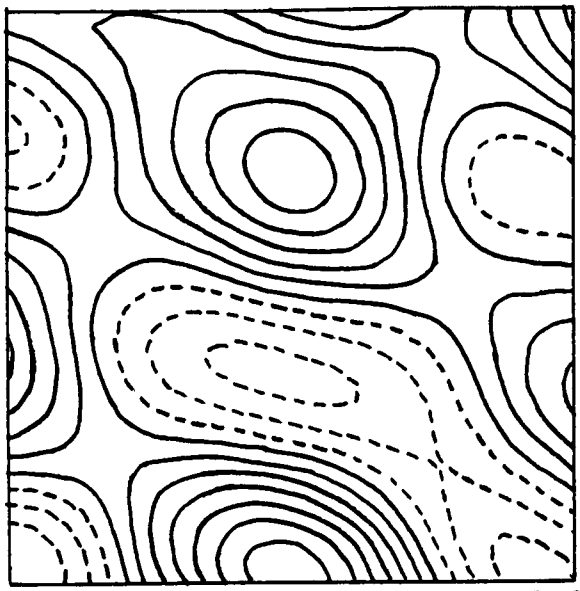


Figure 3.8 The cruising and data collecting path for the statistical forecast experiment that uses data at several time levels. The data are collected at every 4 grid points (62.5 km) along the cruising path.

Figure 3.9 Statistical forecast results using data at several time levels at day

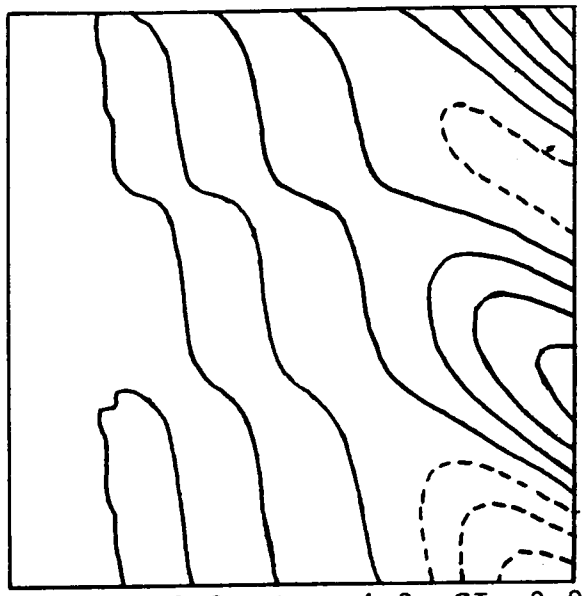
(a) 1 (b) 2 (c) 3
(d) 4 (e) 5 (f) 6
(g) 7 (h) 8 (i) 9
(j) 10 (k) 11 (l) 12
(m) 13 (n) 14 (o) 15
(p) 16 (q) 17 (r) 18
(s) 19.

The figures shown on the left hand side are true fields, while the figures shown on the right hand side are the statistical forecast fields.
The NRMS streamfunction error is shown in (t).

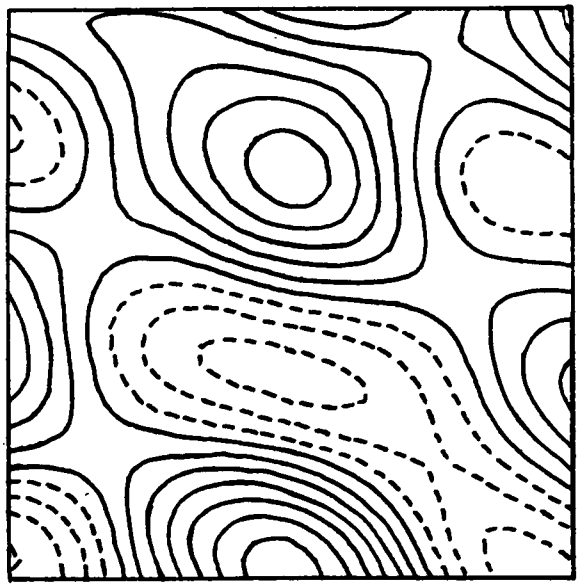


Min= -4.0 Max= 6.6 CI= 1.0

(a)

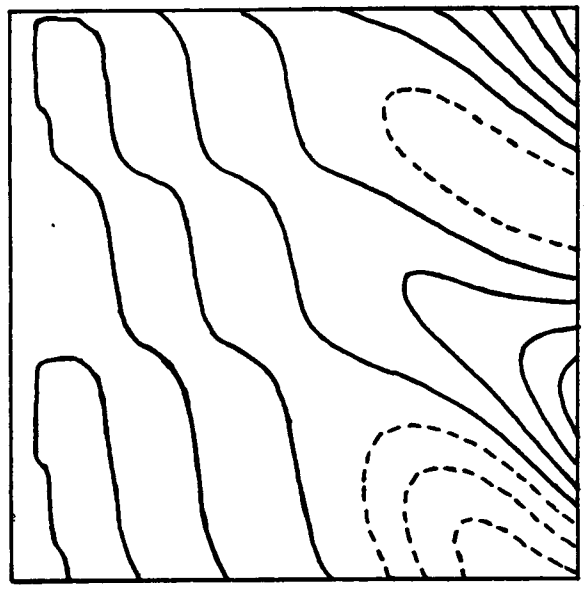


Min= -3.3 Max= 4.3 CI= 0.9

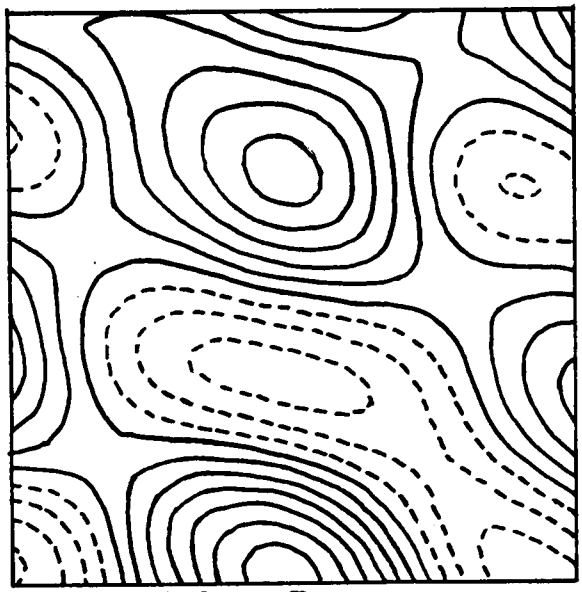


Min= -4.1 Max= 6.6 CI= 1.0

(b)

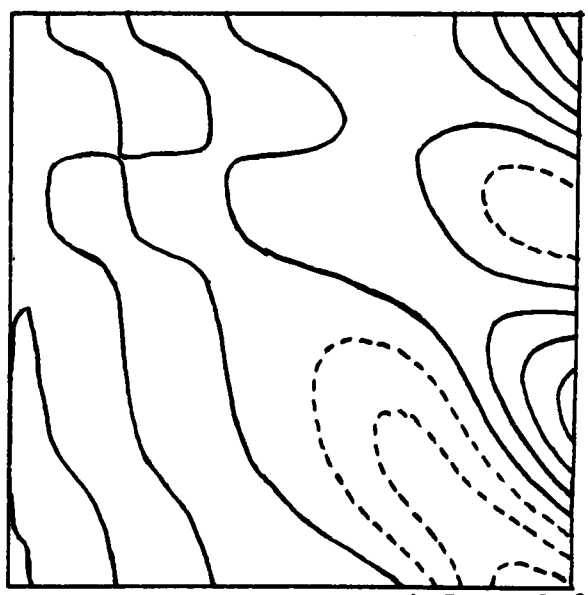


Min= -3.3 Max= 4.5 CI= 0.9



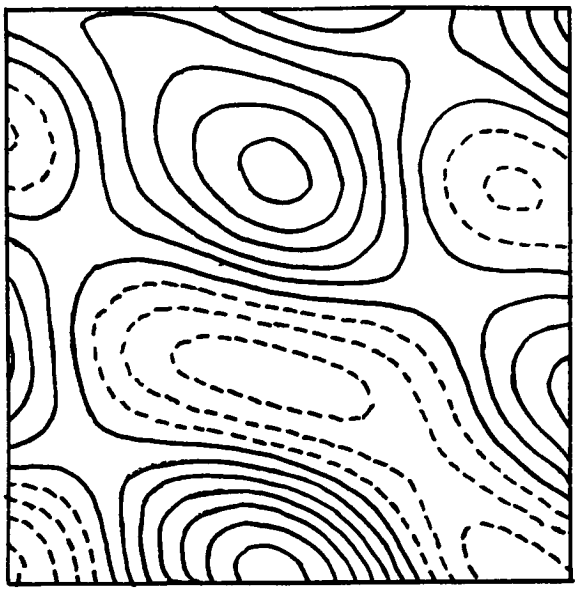
Min= -4.2 Max= 6.5 CI=1.

(c)

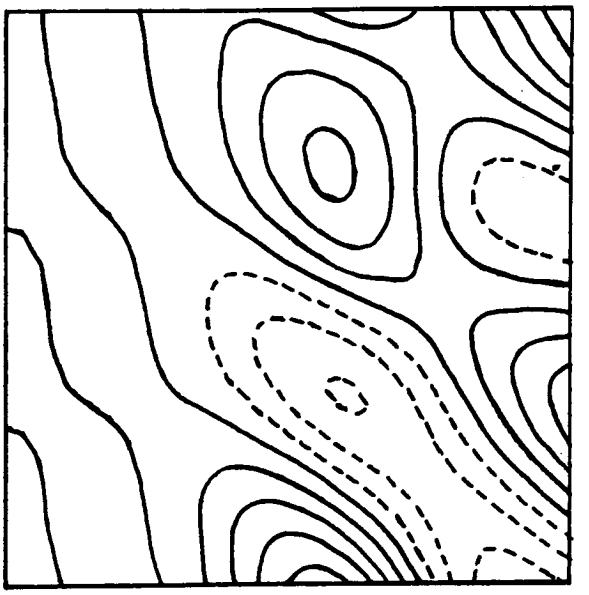


Min=-3.2 Max=4.5 CI=0.9

Figure 3.9 (Cont.)

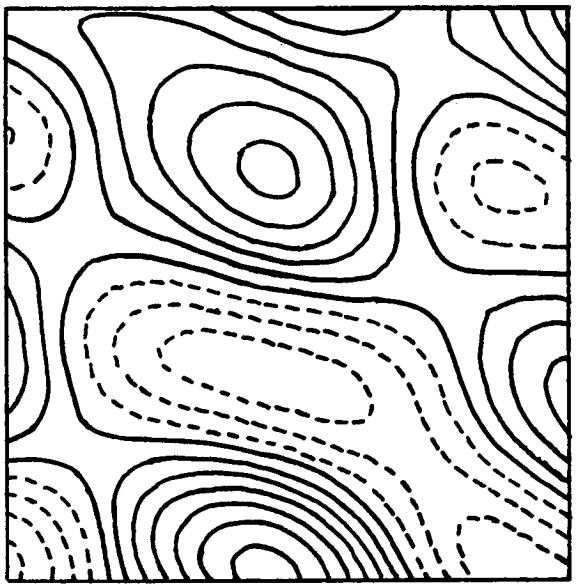


(d)

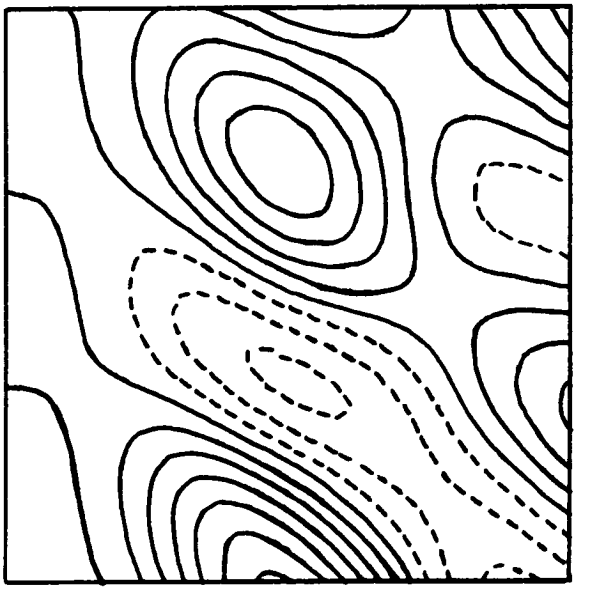


Min=-4.3 Max=6.5 CI=1.0

Min=-3.3 Max=4.5 CI=0.9

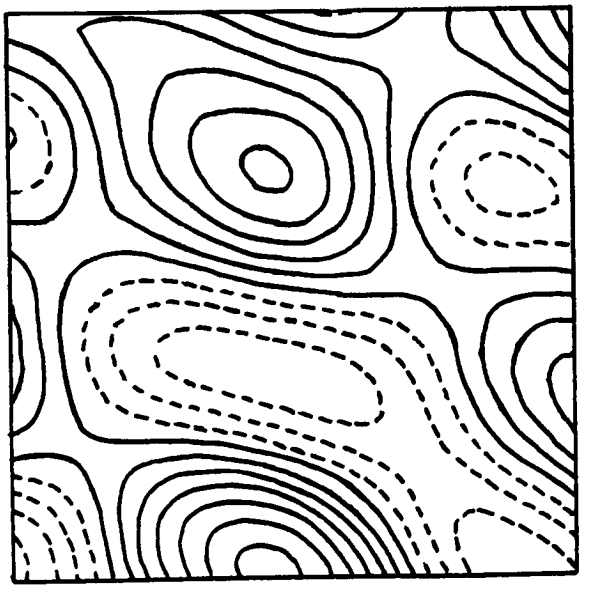


(e)

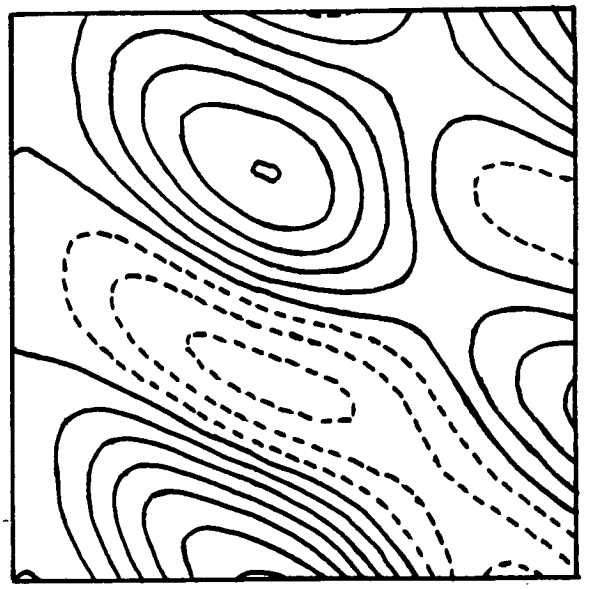


Min=-4.3 Max=6.4 CI=1.0

Min=-3.3 Max=6.2 CI=1.0



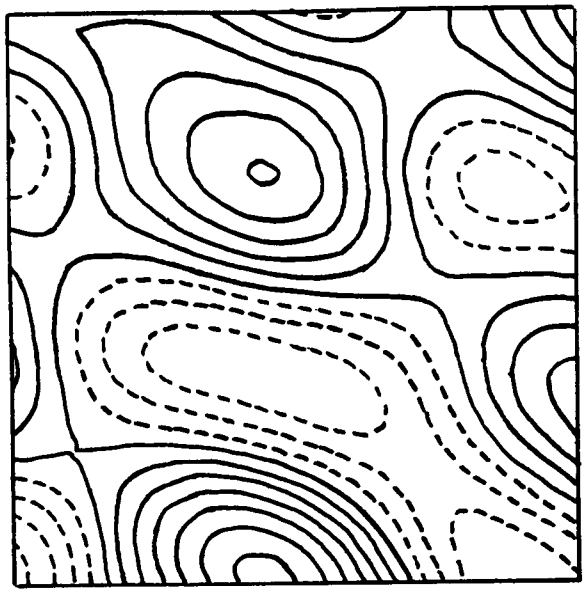
(f)



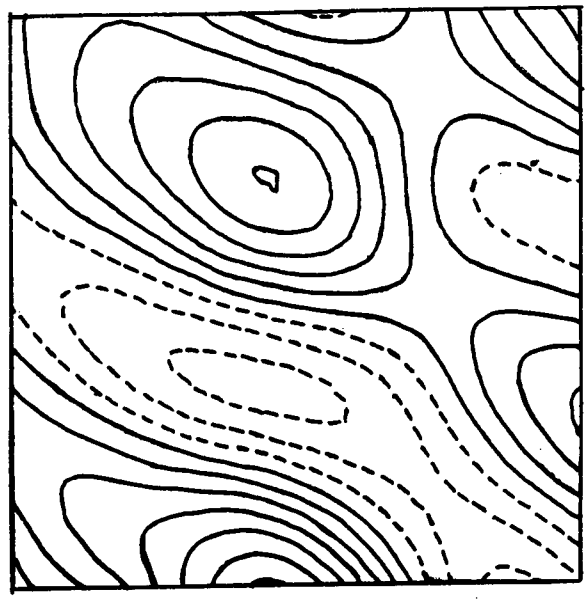
Min=-4.4 Max=6.4 CI=1.0

Figure 3.9 (Cont.)

Min=-3.4 Max=6.2 CI=1.0

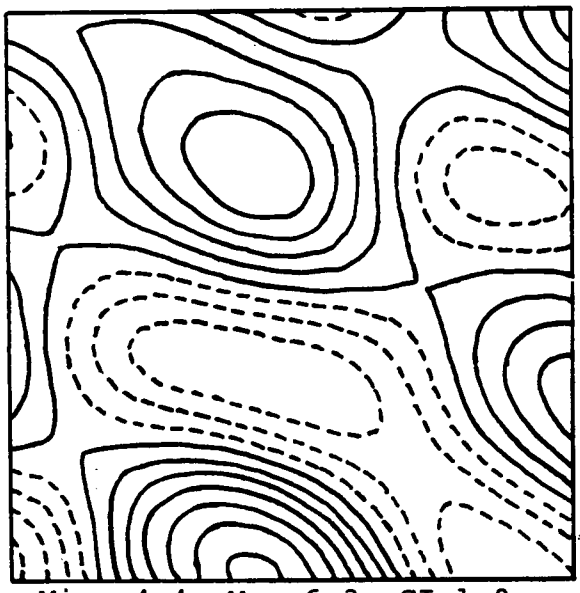


(g)

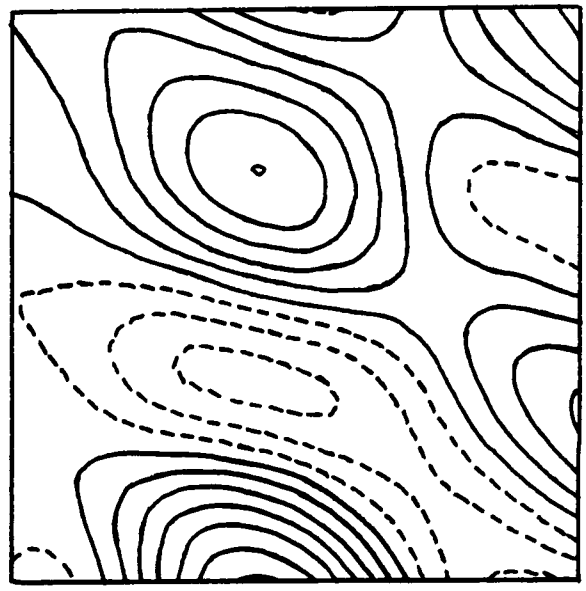


Min=-4.4 Max=6.3 CI=1.0

Min=-3.4 Max=6.1 CI=1.0

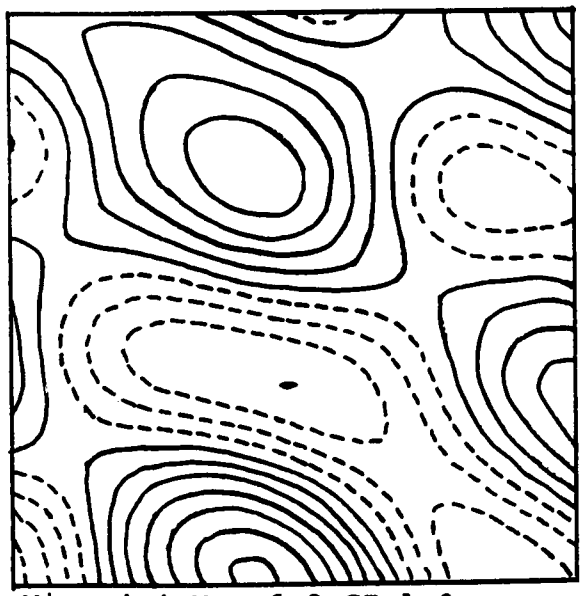


(h)

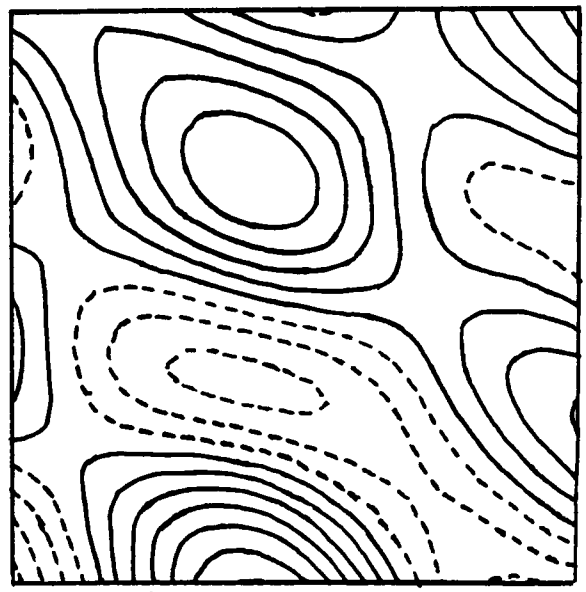


Min=-4.4 Max=6.3 CI=1.0

Min=-3.4 Max=6.2 CI=1.0



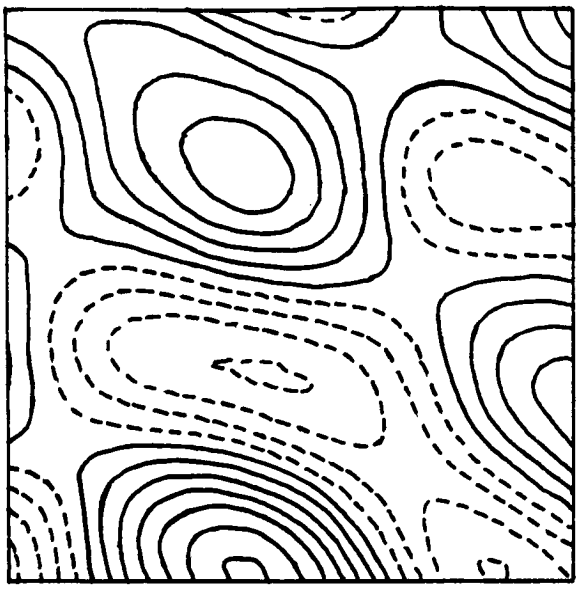
(i)



Min=-4.4 Max=6.2 CI=1.0

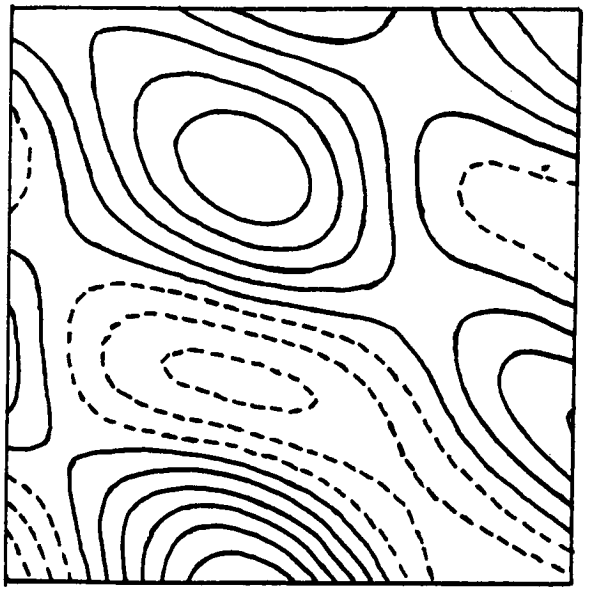
Min=-4.4 Max=6.0 CI=1.0

Figure 3.9 (Cont.)

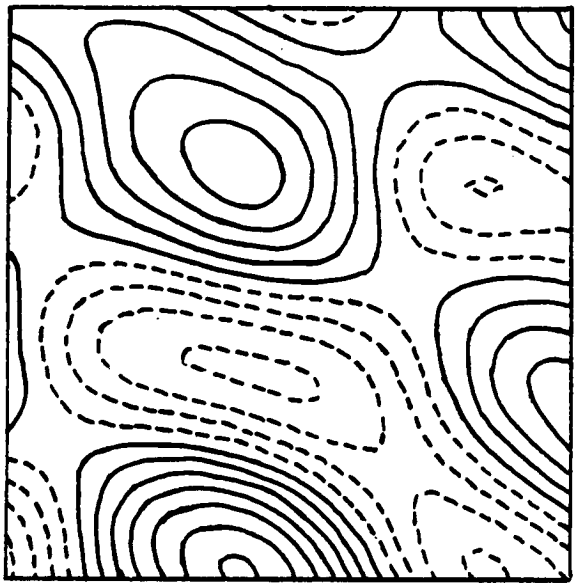


Min=-4.4 Max=6.2 CI=1.0

(j)

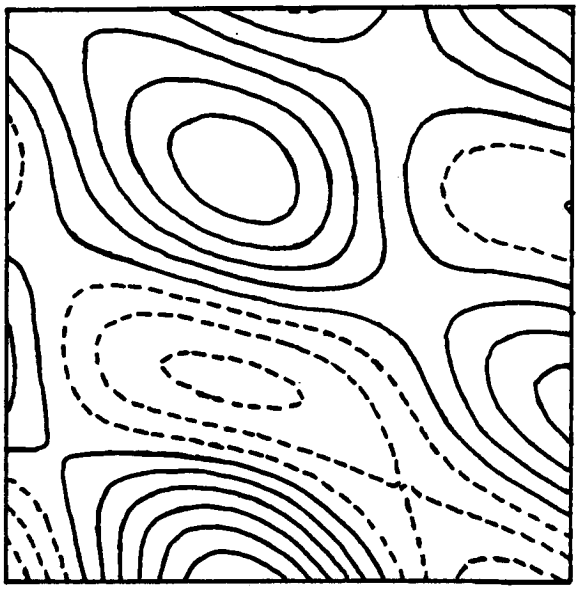


Min=-4.2 Max=6.0 CI=1.0

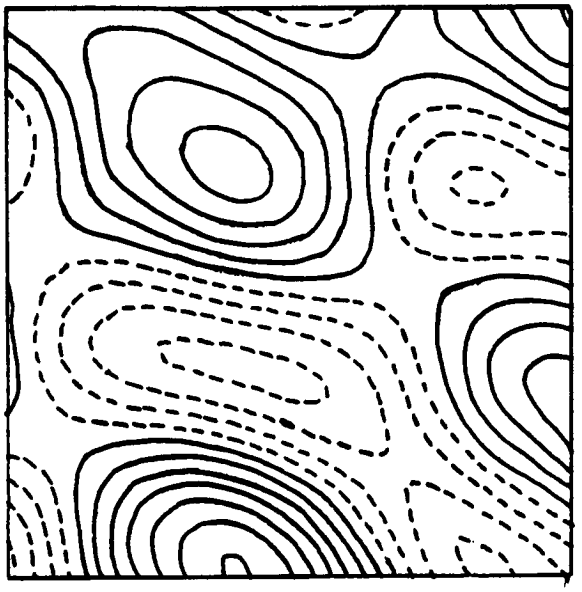


Min=-4.3 Max=6.1 CI=1.0

(k)

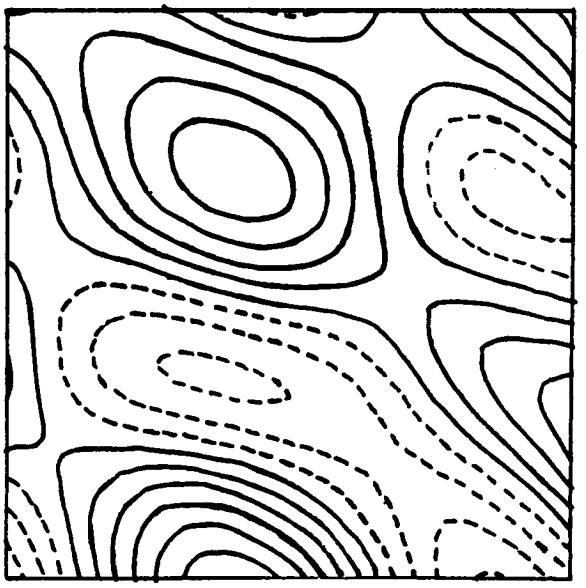


Min=-4.0 Max=5.9 CI=1.0



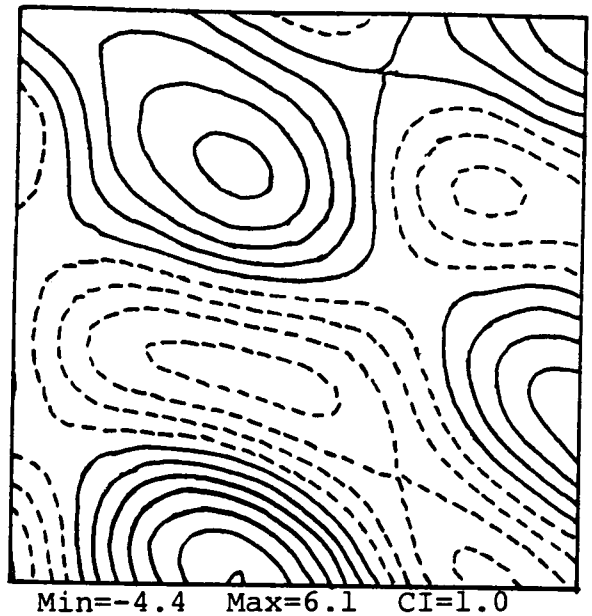
Min=-4.3 Max=6.1 CI=1.0

(l)



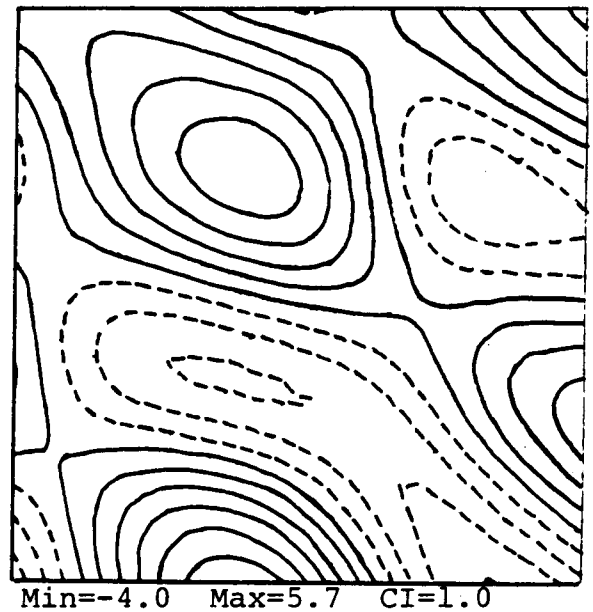
Min=-4.1 Max=5.8 CI=1.0

Figure 3.9 (Cont.)

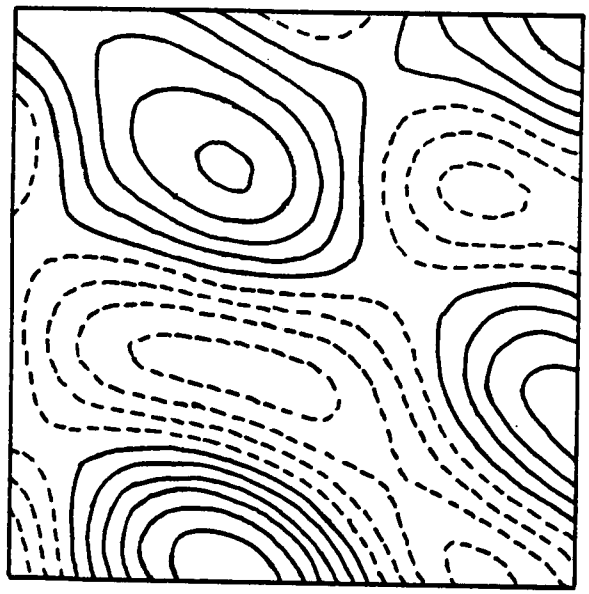


Min=-4.4 Max=6.1 CI=1.0

(m)

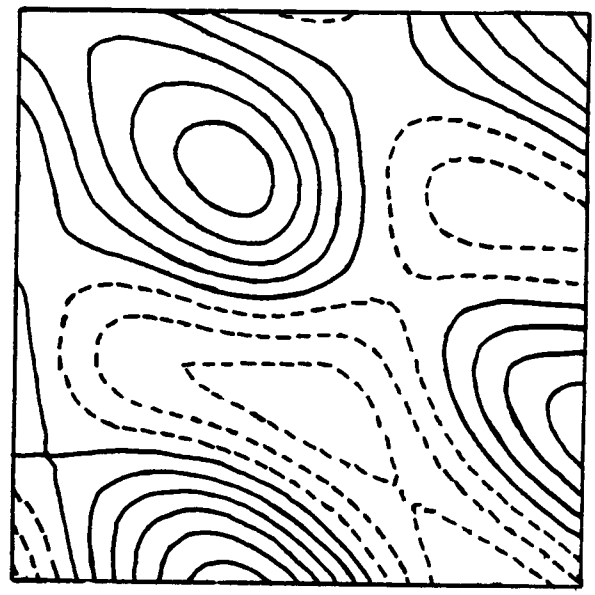


Min=-4.0 Max=5.7 CI=1.0

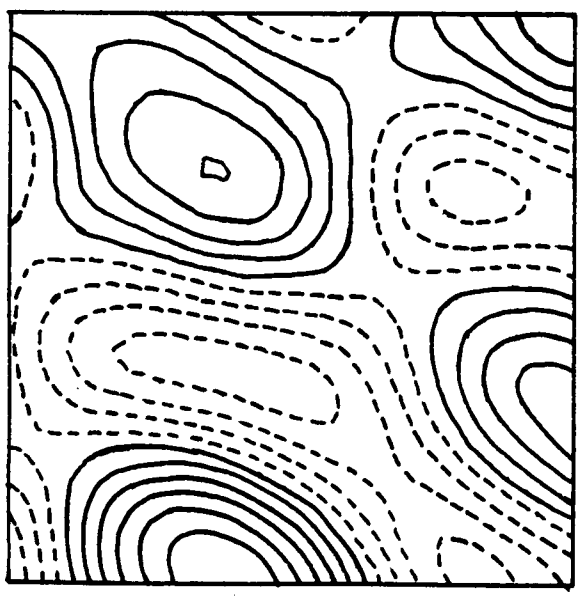


Min=-4.5 Max=6.0 CI=1.0

(n)

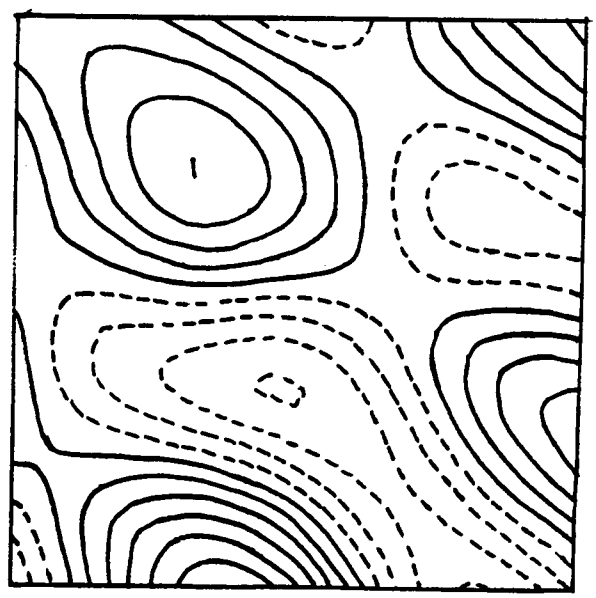


Min=-4.1 Max=5.6 CI=1.0



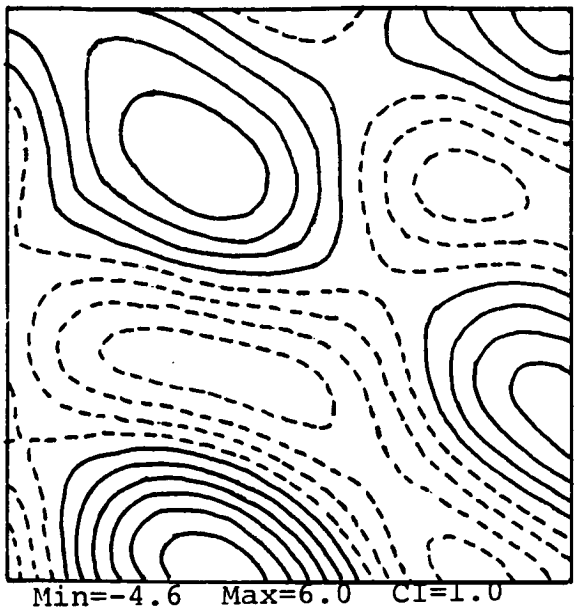
Min=-4.5 Max=6.0 CI=1.0

(o)



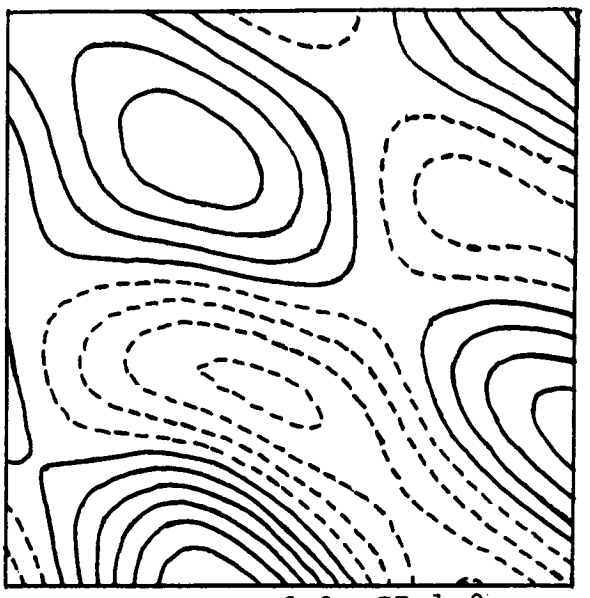
Min=-4.1 Max=5.9 CI=1.0

Figure 3.9 (Cont.)

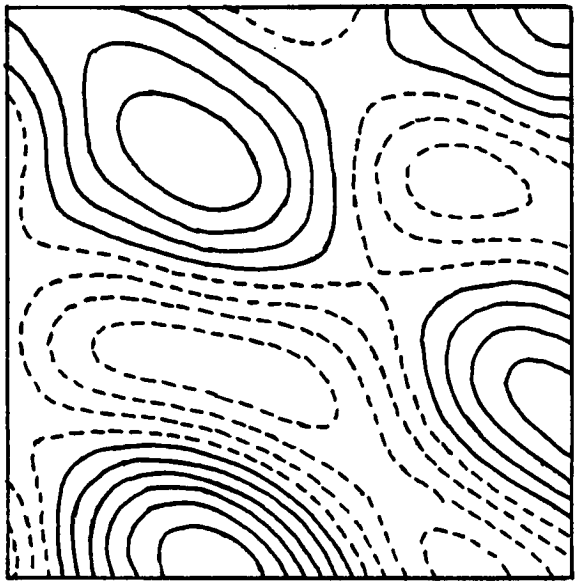


Min=-4.6 Max=6.0 CI=1.0

(p)

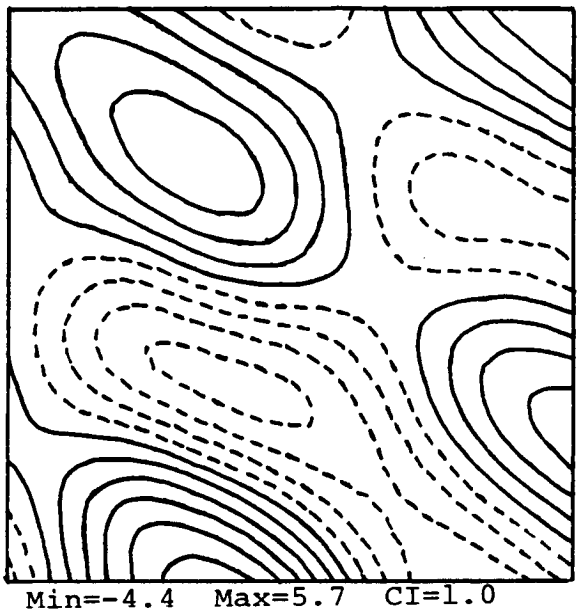


Min=-4.2 Max=6.0 CI=1.0

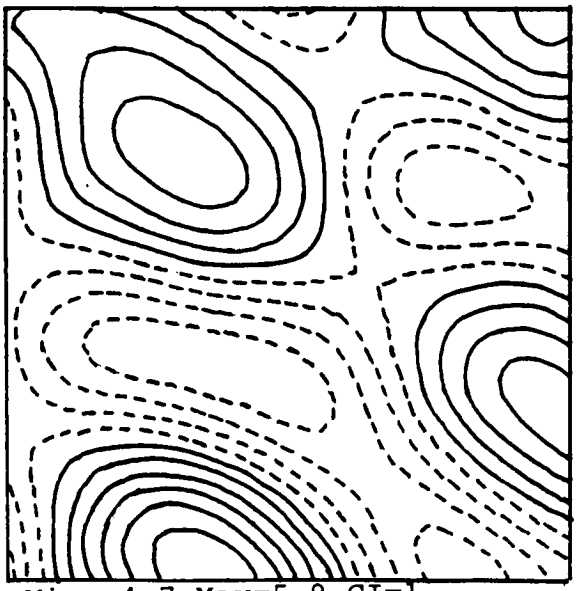


Min=-4.7 Max=5.9 CI=1.0

(q)

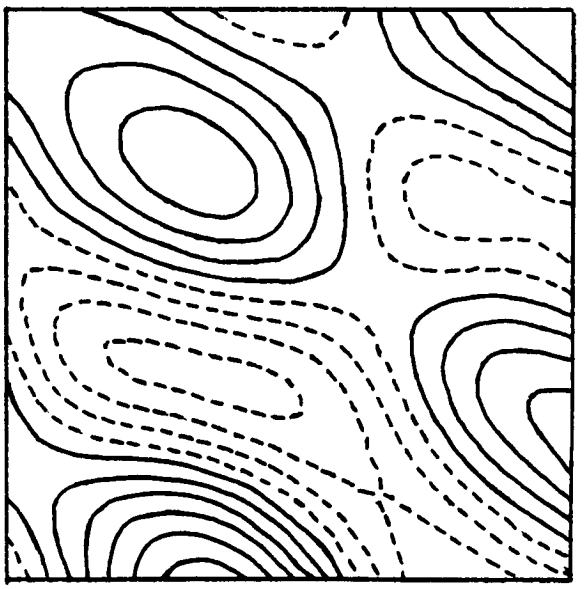


Min=-4.4 Max=5.7 CI=1.0



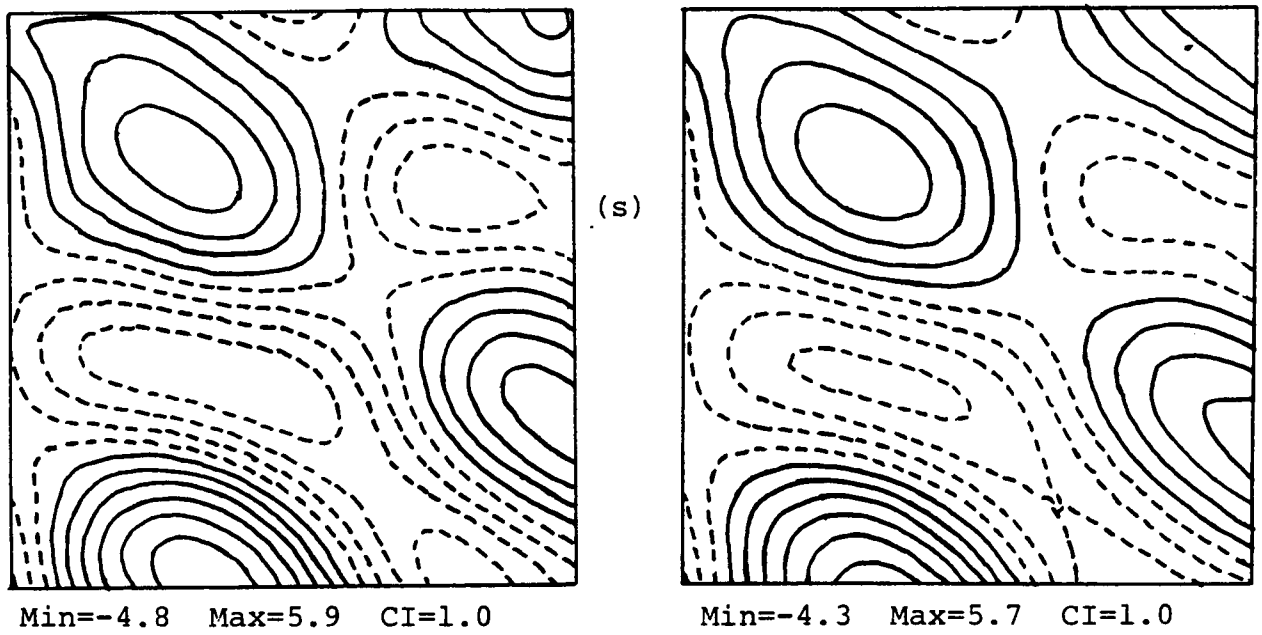
Min=-4.7 Max=5.9 CI=1.

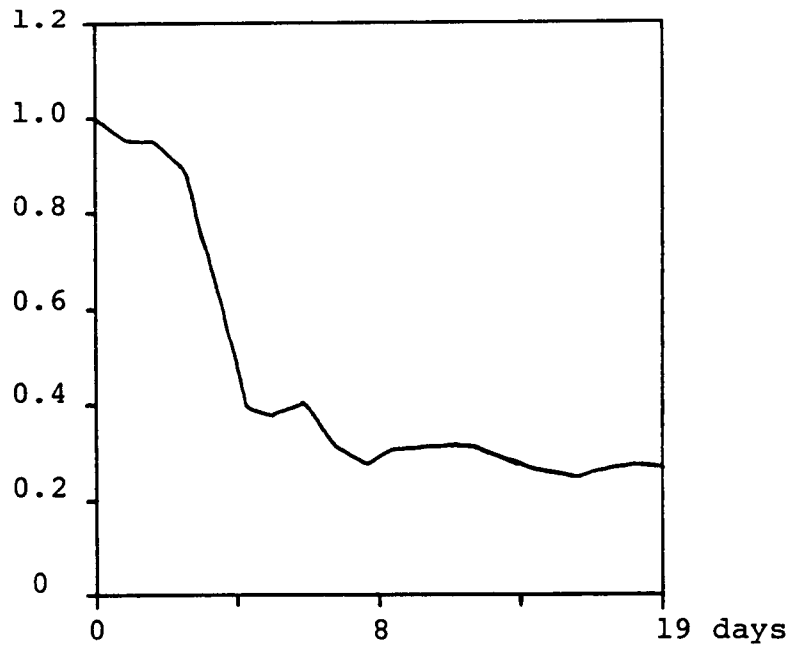
(r)



Min=-4.3 Max=5.7 CI=1.0

Figure 3.9 (Cont.)





(t)

Figure 3.9 (Cont.)

on the left hand side, while the forecast fields are shown on the right hand side. The data used for interpolation on day 1 through day 9 are all the data that have been collected. The result on day 10 is a pure statistical forecast from the previous data because we don't collect any data on this day. The data used for interpolation on days 11 through 19 are the most recent data at all data points, i.e., replace day 1 data by day 11 data, day 2 data by day 12 data, and so forth.

The statistical forecasts are seen to be very good. They produce the qualitative structures of the waves quite well. The parameters used for the statistical forecasts are N = 3, C_{max} = 0.90, Fil = (2,100). The NRMS streamfunction error is shown on Figure 3.9(t). It can bee seen that after one lap of cruising, the error is maintained at about 30%, which is about the level of error expected for statistical forecasts with parameters S = 4, Δt = 4 (this is the average of the time steps that the data are forecasted forward in time).

CHAPTER IV
THE COMBINED STATISTICAL AND DYNAMICAL APPROACH

IV.1 Optimal combination of two estimates

In this section we introduce another statistical method which is used to obtain a better estimate from several a priori estimates of the same quantity. Let ψ^1, ψ^2 be two a priori estimates of a random variable ψ with the following characteristics:

$$\overline{(\psi^1-\psi)} = B_1 \tag{4.1}$$

$$\overline{(\psi^2-\psi)} = B_2 \tag{4.2}$$

$$\overline{(\psi^1-\psi)^2} = E_1^2+B_1^2 \tag{4.3}$$

$$\overline{(\psi^2-\psi)^2} = E_2^2+B_2^2 \tag{4.4}$$

$$\overline{(\psi^1-\psi)(\psi^2-\psi)} = \gamma E_1E_2+B_1B_2 \tag{4.5}$$

where the overbar denotes ensemble average.

Let ψ^e be a convex linear combination of ψ^1-B_1 and ψ^2-B_2, i.e.

$$\psi^e = \alpha(\psi^1-B_1)+(1-\alpha)(\psi^2-B_2) \qquad 0\le\alpha\le 1 \tag{4.6}$$

We are interested in determining the optimal ψ^e that minimizes the mean square error $E^2 = \overline{(\psi^e-\psi)^2}$.

It is straightforward to show that the optimal estimation ψ^{e^*} has weighting coefficients

$$\alpha^* = \frac{E_2^2-\gamma E_1E_2}{E_1^2+E_2^2-2\gamma E_1E_2} \tag{4.7}$$

$$1-\alpha^* = \frac{E_1^2-\gamma E_1E_2}{E_1^2+E_2^2-2\gamma E_1E_2} \tag{4.8}$$

and mean square estimation error

$$E^{*2} = \frac{E_1^2 E_2^2 (1-\gamma^2)}{E_1^2 + E_2^2 - 2\gamma E_1 E_2} \tag{4.9}$$

Figures 4.1 and 4.2 show E^{*2} and $(\alpha^*, 1-\alpha^*)$ as a function of the correlation coefficient γ respectively.

Here we assume the a priori estimates are biased and correlated. Of course, we could have removed the biases if they had been already known, but sometimes we don't know the biases exactly. For example, a dynamical forecast with persistent boundary conditions gives biased estimates of the field at a later time with uncertain biases. In this case, the biases have to be estimated. The present formulations are adopted so that they can be easily applied.

The formula for optimal combination of estimates are very useful for our forecasting studies here, because sometimes it is possible to obtain more than one estimate of the field--e.g., via dynamical forecast and statistical forecast. The optimal combination gives us a better estimate of the field than each individual one.

The optimal combination of estimates can be easily extended to more than two a priori estimates. This has special implications in ocean modellings because of the availability of a wide variety of data sources. If we form an a priori estimate from the data of each source, then we can use the optimal combination formula to obtain a better estimate of the synoptic field and finally assimilate this information into the model. Some ex-

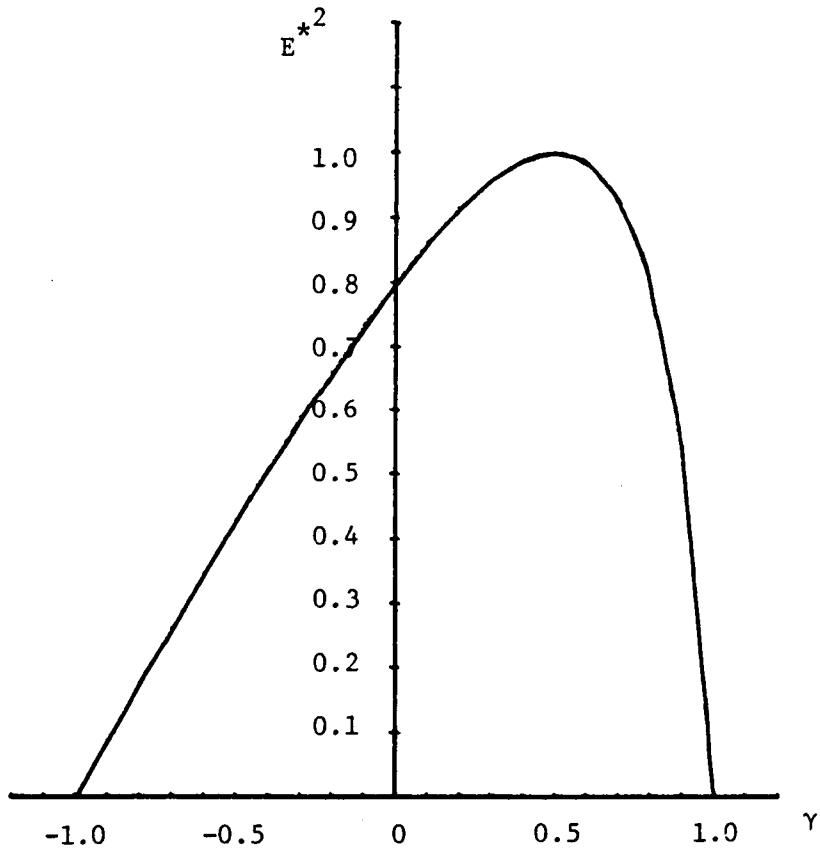


Figure 4.1 The minimal mean square error E^{*2} in combining two a priori estimates as a function of the correlation coefficient γ($E_1^2=1$, $E_2^2=4$).

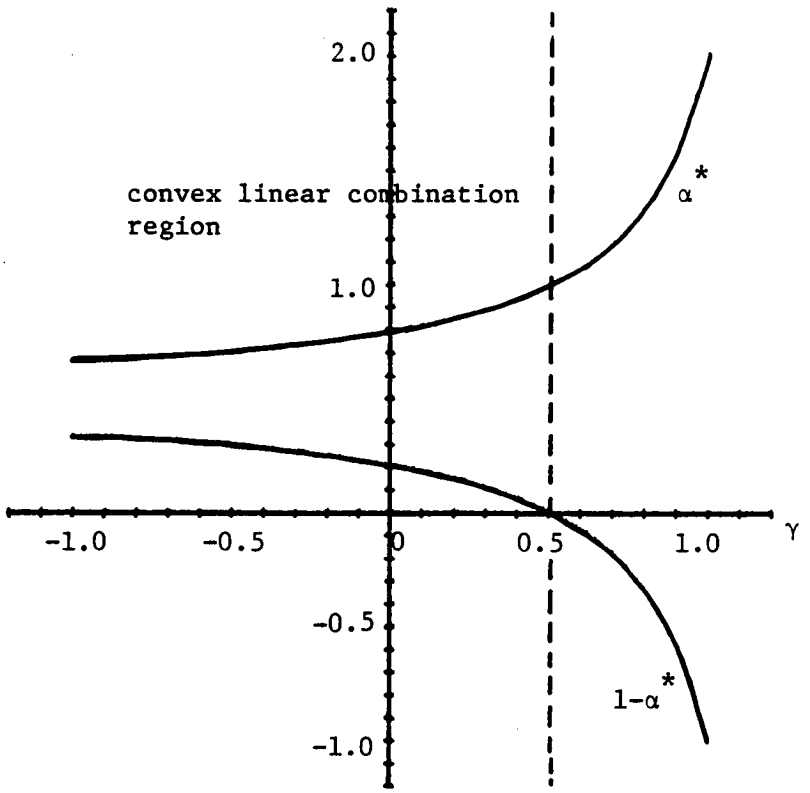


Figure 4.2 The optimal weighting coefficients in combining two a priori estimates as a function of the correlation coefficients γ ($E_1^2=1$, $E_2^2=4$). The region with α^*, $(1-\alpha^*)\geq 0$ is the convex linear combination region.

amples of applying the optimal combination will be given in Section IV.3.

IV.2 The Error Model

If we are going to combine two a priori estimated fields together, from Section IV.1 we have to know five statistical parameters, viz., B_1, B_2, E_1^2, E_2^2 and γ at each grid point (in general $2n + \frac{n(n-1)}{2}$ parameters in combining n a priori estimates). In the case that a field has one thousand grid points, we have to know five thousand statistical parameters in order to combine two fields together. This is a formidable quantity of information to obtain or estimate. Furthermore, if we are going to drive the forecasting model forward in time, these statistics have to be known at several time levels. In some special cases, e.g., in a closed domain with linear system dynamics, the evolution of these statistics can be computed analytically from the initial statistics, the system dynamics and noise parameters. But in an open domain with nonlinear system dynamics, to compute the evolution of these statistics analytically is totally out of the question.

The approach we adopt here is to compute these statistics directly from the estimated fields. This is possible due to the availability of the simulation verification data which is provided by the exterior calculation described in Chapter II. However, we do not compute these statistics on a grid point by grid point basis. We group those grid points which are likely

to have the same statistics together by looking roughly at the error map. Divide the whole domain into regions which are groups of grid points. The statistics in each region are then treated as uniform and are computed from the simulated error data. An example of constructing this error model is shown in Figure 4.3.

In the beginning, we use these actual computed statistics to combine two given field estimates together. This is to provide an optimal estimate for a particular error model chosen. Later, these statistics are replaced by estimated values which are the averages of the actual computed statistics from several realizations. We use these estimated statistics to run the same experiment again. If the result is not too far different from the previous one, these estimated statistics will be accepted. It is these estimated statistics that are intended to be used in the forecasting studies which employ real ocean data as discussed in Chap. V.

IV.3 Examples of the Combined Statistical and Dynamical Approach

In this Section we explore various ways of employing statistical methods in the dynamical model. The statistical methods we use are essentially the space-time objective analysis introduced in Chapter III and the optimal combination of estimates introduced in Section IV.1.

The first example uses the objective analysis to map the coarsely distributed data onto the computational grid in order to provide the initial and boundary conditions for the dynamical model. The second example uses the space-time objective

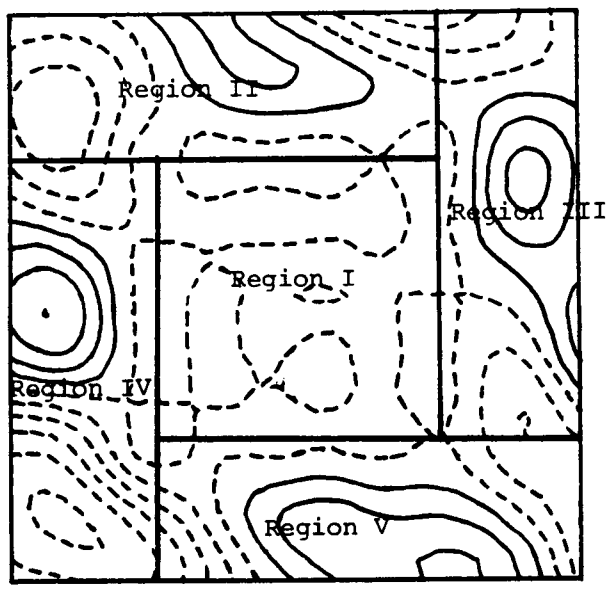


Figure 4.3 An example showing the construction of an error model from the error map.

analysis to forecast the boundary condition for the dynamical model either from an initial condition or from the previous dynamical forecast field. The third example applies the optimal combination of estimates formulas to assimilate a patch of recently acquired data into the dynamical model. The last example applies the optimal combination of estimates formulas to obtain a better estimate of the field from the persistent statistical forecast and persistent dynamical forecast.

(i) Initial and boundary conditions sampling experiments*

In the first example, we perform several simulation experiments using various kinds of data sampling schemes to provide the initial and boundary conditions for the dynamical model. The streamfunction data are collected in the domain at an initial time and near the boundary at all future times. They are not collected at every grid point. Instead, the objective analysis is used to map these coarsely distributed data onto the computational grid. Some experiments use true initial and/or boundary vorticity, while others derive the initial and/or boundary vorticity from the streamfunction by using the second order Laplacian scheme, after the streamfunction data have been wholly interpolated. All the experiments performed in this regard are summarized in Table 4.1.

---

* Private communication by A.R. Robinson and J. Groisser.

Experiment No.	Initial Condition		Boundary Condition	
	ψ	ζ	ψ	ζ
4.1	"2"	D	V	V
4.2	"3"	D	V	V
4.3	"4"	D	V	V
4.4	V	V	"X"	D
4.5	V	V	"V"	D
4.6	V	V	"BX"	D
4.7	V	V	"BV"	D
4.8	V	V	"N"	D
4.9	"4"	D	"BV"	D
4.10	"4"	D	"BVP5"	D

Table 4.1 Summary of initial and boundary conditions sampling experiments.

where

"2" denotes that the data are sampled at every 2 grid points (31.25 km).

"3" denotes that the data are sampled at every 3 grid points (46.875 km).

"4" denotes that the data are sampled at every 4 grid points (62.5 km).

V denotes that the true (veritas) data are used or the data are sampled at every 1 grid point (15.625 km).

D denotes that the vorticity is derived from the streamfunction using the second order Laplacian scheme.

"X" denotes that the boundary data are sampled using the "X" scheme.

"V" denotes that the boundary data are sampled using the "V" scheme.

"BX" denotes that the boundary data are sampled using the "BX" scheme.

"BV" denotes that the boundary data are sampled using the "BV" scheme.

"N" denotes that the boundary data are sampled using the "N" scheme.

"BVP5" denotes that the boundary data are sampled using the "BV" scheme and kept persistent for 5 time steps.

The boundary data sampling schemes "X", "V", "BX", "BV", "N" are show in Figure 4.4. The "X" scheme samples 192 points out of 384 in the boundary strip; the "V" scheme samples 128 points out of those 384; the "BX" scheme, 96 points out of 384; the "BV" scheme samples 64 points, and the "N" scheme, 76 points, both out of 384 in the boundary strip.

The NRMS streamfunction errors for experiments 4.1-4.10 are shown in Figure 4.5. They are compared with that of experiment 4.11, which is the P5 experiment. From this figure, we observe the following:

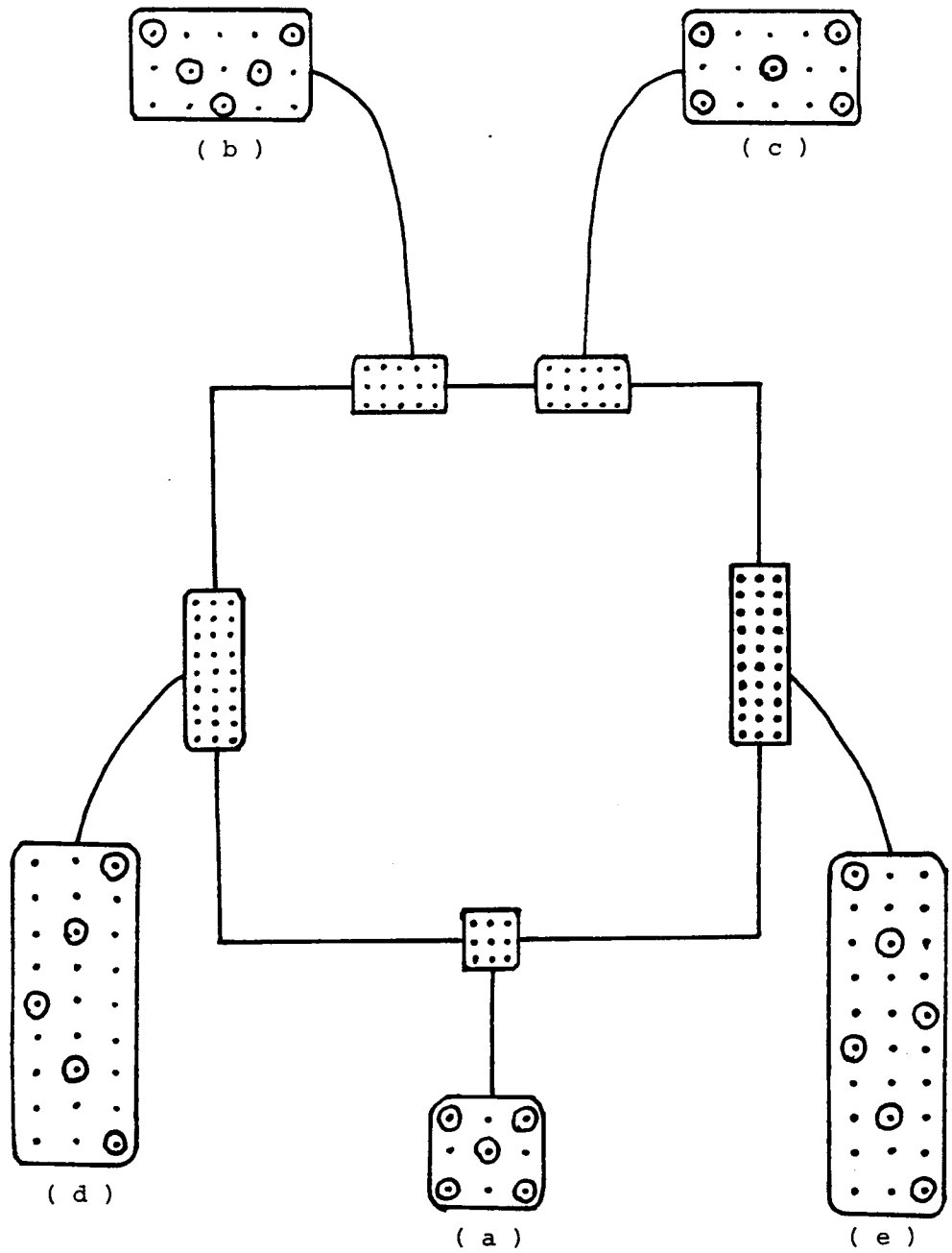


Figure 4.4 Different boundary data sampling schemes.
(a) "X" : samples 192/384 in the boundary strip.
(b) "V" : samples 128/384 in the boundary strip.
(c) "BX": samples 96/384 in the boundary strip.
(d) "BV": samples 64/384 in the boundary strip.
(e) "N" : samples 76/384 in the boundary strip.

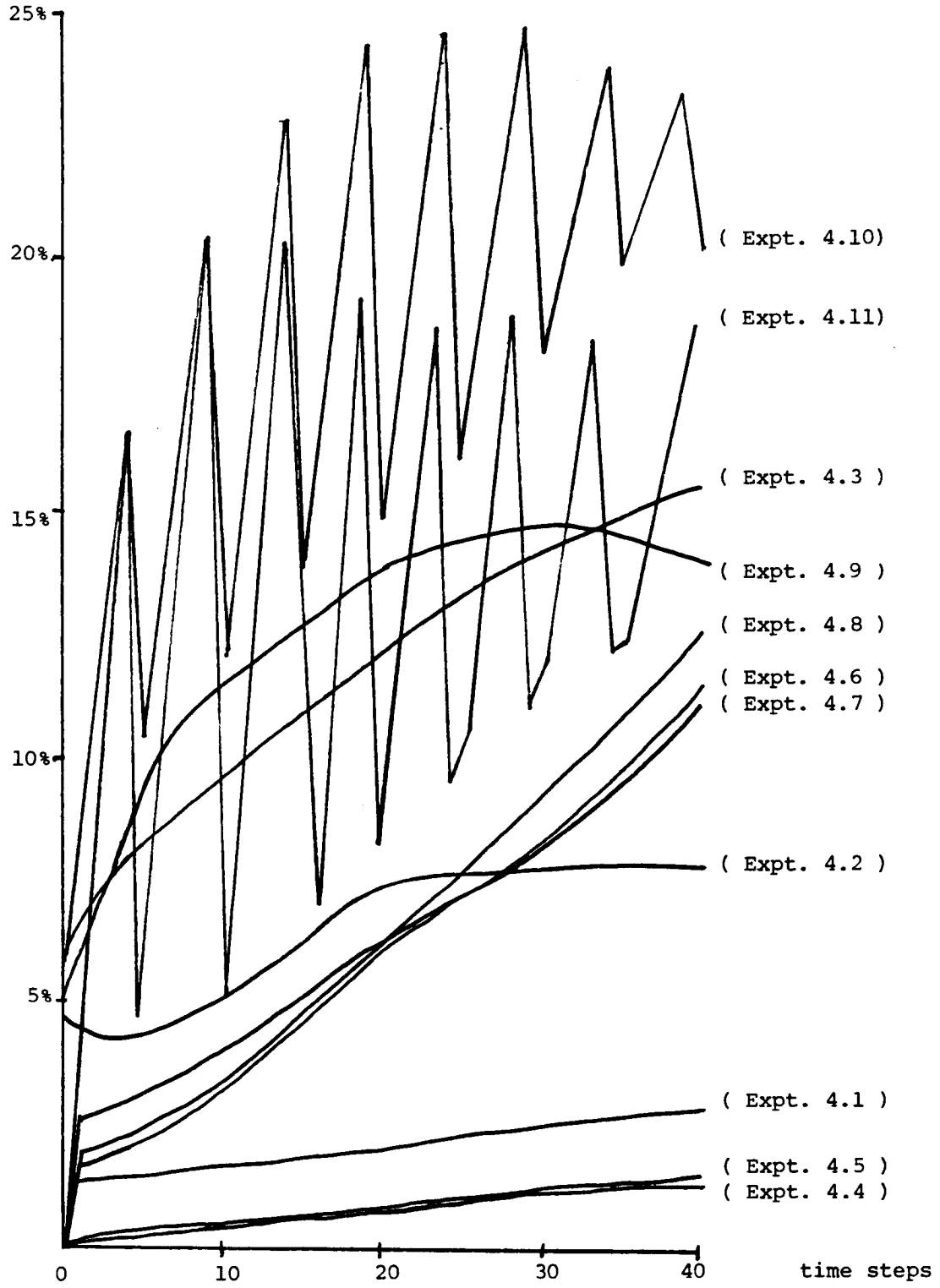


Figure 4.5 NRMS streamfunction errors of the initial and boundary condition data sampling experiments.

(a) If the boundary condition sampling scheme is kept the same, the NRMS error increases as the initial condition sampling scheme becomes coarser.

(b) Boundary condition sampling schemes "X" and "V" produce almost the same results. Boundary condition sampling schemes "BX", "BV", "N" produce almost the same result. The NRMS error, in general, increases as the number of data points sampled in the boundary strip decreases -- at least in the beginning.

(c) Although we have, in experiment 4.9, reduced the initial condition sampling to only 1/16 of the interior points, boundary condition sampling to only 1/6 of the boundary strip, the NRMS error is still below 15%.

(d) The NRMS errors from different error sources are not independent. The error caused by coarse sampling in both initial and boundary conditions is not separable into errors caused by coarse sampling in initial condition and boundary condition individually. Experiments 4.3, 4.7, 4.9 illustrate this remark.

(ii) Dynamical Forecast using Statistical Forecast Boundary Condition

In a closed domain, the boundary condition is time invariant. It is always perfect and, theoretically, there is no contribution to the forecast error from the boundary condition. In an open domain, the boundary condition changes with time. It is not known beforehand. This uncertainty in the boundary

condition will deteriorate the dynamical forecast. Furthermore, from Section II.5, we know that the boundary condition is the principle factor in determining the forecast error. Thus, to obtain a better dynamical forecast, we first have to obtain a better estimate of the boundary condition.

In this example, we are given a perfect initial field and are asked to forecast the field forward in time as accurately as possible. Here we try the statistical forecast model (space-time objective analysis) to predict the boundary condition, and then use the predicted boundary condition to drive the dynamical forecast model. Two experiments are performed here in predicting the boundary condition. The first experiment always uses the perfect initial field as the observation data to forecast, in the statistical model, boundary conditions at later times. The second experiment uses the previous dynamical forecast field as the observation data to forecast, again in the statistical model, the boundary condition at the next time step. Of course some other strategies are possible. The NRMS streamfunction errors of these two experiments are shown in Figure 4.6. They are compared with the error in the persistent experiment, in which the boundary condition is kept frozen from the initial time as introduced in Section II.4. Curves (a), (b) and (c) are for the persistent dynamical forecast, dynamical forecast using the statistical forecast boundary condition from the perfect initial field, dynamical forecast using the statistical forecast boundary condition from the

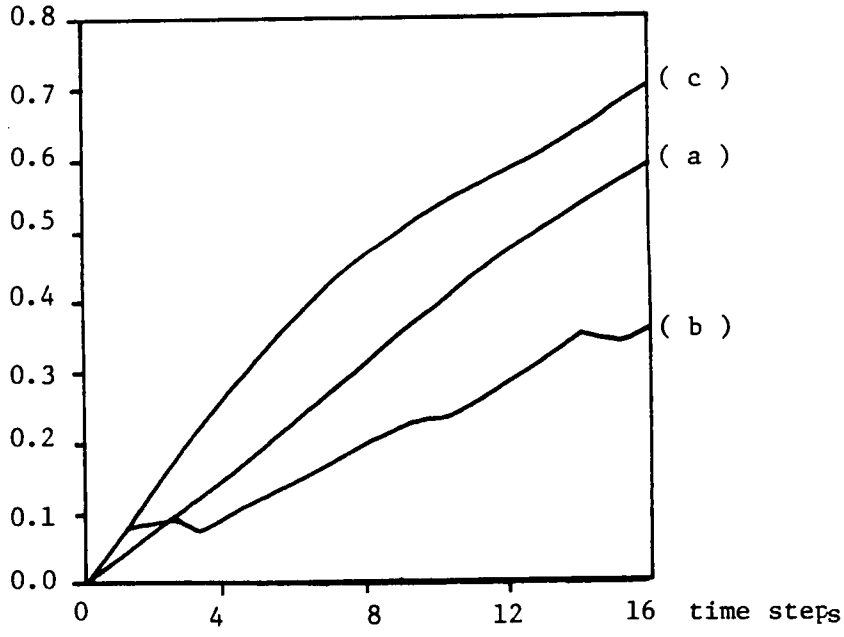


Figure 4.6 The NRMS streamfunction errors of the
(a) persistent dynamical forecast
(b) dynamical forecast using the statistical forecast boundary condition from the initial field.
(c) dynamical forecast using the statistical forecast boundary condition from the previous dynamical field.

previous dynamical forecast field respectively.

The calculations are performed for 16 time steps (about 2 weeks). In the beginning, the forecasts using the statistical forecast boundary condition are not better than the persistent dynamical forecast. This is because the statistical model is not able to resolve the wave speeds for the grid interval Δx and time step Δt chosen. The waves propagate about 1/3 grid interval per time step. When it searches for the most correlated point at the first time step, in the statistical model, it will assume the wave is stagnant or has moved one grid interval depending on the exact values of the correlation function for $\Delta t = 1$ (time step). Thus, the statistical forecast boundary conditions for small Δt are not good, hence the dynamical forecast using these boundary conditions. However, for large Δt, the statistical forecast boundary conditions become better estimates than the persistent boundary conditions. Therefore, the dynamical forecast using these boundary conditions improves. This can be seen from curve (b) which has significant improvement over curve (a) for large Δt (33% vs. 58% at the end of 16 time steps). Curve (c) is worse than curve (a) because the boundary condition is always forecasted one time step forward from the previous dynamical field and the statistical forecast for one time step can not produce good results as explained earlier. The improvement of curve (b) over (a) is gained by using a better estimate of the boundary condition, which can be used as a measure of the value of the better

boundary information. This example tells us that, given only an initial field, the persistent dynamical forecast is not the best we can do. Obtaining a better estimate of the boundary condition by some scheme will improve the dynamical forecast. The statistical forecast of the boundary condition is one, but there are still better; in Section IV.4, we will see some other schemes.

(iii) Patch Updating in the Interior

In Section II.5, when we updated the interior, we updated the whole field. A whole field of data is a large quantity of information. Collecting it would take up valuable resources. Moreover, only partial data, sparsely scattered through the region, are available--especially in the oceanic case. Yet these data, no matter how few, will give us new information about the system. With them we can find a better estimate of the dynamic field. There are different ways of assimilating these new data into the dynamical model. They can be either inserted directly into the dynamical forecast field or else interpolated into the whole region and then combined with the dynamical forecast field, etc. For discussions of various data assimilation techniques, see Bengtsson (2, 3, 4) and Ghil (11).

The data assimilation method, which is used here, interpolates the sparse data into the whole region by the objective analysis formula introduced in Section III.1, and then com-

bines the interpolated field with the dynamical forecast field by a formula (introduced in Section IV.1) for the optimal combination of estimates. This approach is slightly different from the one commonly practiced in Meteorology, which interpolates the difference between the observation and the forecast to find a correction term to the forecast field (Rutherford [27]). It uses the same interpolation technique, but the statistics required are different. The differences between these two approaches are explained in detail below.

Let ψ^t be the true field

ψ^f be the dynamical forecast field

ψ^o be a set of observations at location x_i, i=1, ..N

ψ^a be the interpolated field

ψ^e be the optimal estimate field

assume

$$\psi^o(x_i) = \psi^t(x_i) \qquad (i = 1, \cdots N) \tag{4.10}$$

(i.e., the measurements are error free)

Our approach is first to obtain $\psi^a(x)$ from ψ^o by objective analysis

$$\psi^a(x) = \overline{\psi^t} + \sum_{i=1}^{N} \alpha_i(x)\,(\psi^o(x_i) - \overline{\psi^t}) \tag{4.11}$$

where

$$\alpha_i(x) = \sum_{j=1}^{N} C_{ij}^{-1}\, \overline{(\psi^t(x) - \overline{\psi^t})(\psi^t(x_j) - \overline{\psi^t})} \tag{4.12}$$

$$C_{ij} = \overline{(\psi^t(x_i) - \overline{\psi^t})(\psi^t(x_j) - \overline{\psi^t})} \tag{4.13}$$

then we combine $\psi^I(x)$ and $\psi^f(x)$ to obtain $\psi^e(x)$

$$\psi^e(x) = \alpha\,(\psi^I(x) - B_I) + (1-\alpha)(\psi^f(x) - B_f) \tag{4.14}$$

where

$$B_I = \overline{\psi^I} - \overline{\psi^t} \tag{4.15}$$

$$B_f = \overline{\psi^f} - \overline{\psi^t} \tag{4.16}$$

$$\alpha = \frac{E_f^2 - \gamma E_I E_f}{E_I^2 + E_f^2 - 2\gamma E_I E_f} \tag{4.17}$$

$$1-\alpha = \frac{E_I^2 - \gamma E_I E_f}{E_I^2 + E_f^2 - 2\gamma E_I E_f} \tag{4.18}$$

$$E_I^2 = \overline{(\psi^I - \psi^t)^2} - B_I^2 \tag{4.19}$$

$$E_f^2 = \overline{(\psi^f - \psi^t)^2} - B_f^2 \tag{4.20}$$

$$\gamma = \frac{\overline{(\psi^I - \psi^t)(\psi^f - \psi^t)} - B_I B_f}{E_I E_f} \tag{4.21}$$

Thus the statistics required in our approach are

$$\overline{\psi^t} \tag{4.22}$$

$$\overline{\psi^I} \tag{4.23}$$

$$\overline{\psi^f} \tag{4.23}$$

$$\overline{\psi^t(x)\,\psi^t(y)} \tag{4.25}$$

$$\overline{(\psi^I - \psi^t)^2} \tag{4.26}$$

$$\overline{(\psi^f - \psi^t)^2} \tag{4.27}$$

$$\overline{(\psi^I - \psi^t)(\psi^f - \psi^t)} \tag{4.28}$$

and the data are assimilated in two steps.

The approach commonly practiced in Meteorology as described in Rutherford (27) is to interpolate the difference between the observation and the forecast to find a correction term to the forecast field

$$\psi^e(x) = \bar{\psi}^t + (\psi^f(x) - \bar{\psi}^f) + \sum_{i=1}^{N} \beta_i (\psi^o(x_i) - \psi^f(x_i) - \bar{\psi}^o + \bar{\psi}^f) \qquad (4.29)$$

where

$$\beta_i = D_{ij}^{-1} \overline{[(\psi^t(x) - \bar{\psi}^t - \psi^f(x) + \bar{\psi}^f)(\psi^t(x_j) - \bar{\psi}^t - \psi^f(x_j) + \bar{\psi}^f)}} \qquad (4.30)$$

$$D_{ij} = \overline{(\psi^t(x_i) - \bar{\psi}^t - \psi^f(x_i) + \bar{\psi}^f)(\psi^t(x_j) - \bar{\psi}^t - \psi^f(x_j) + \bar{\psi}^f)} \qquad (4.31)$$

The statistics required in this approach are

$$\bar{\psi}^t \qquad (4.32)$$

$$\bar{\psi}^f \qquad (4.33)$$

$$\overline{(\psi^t(x) - \psi^f(x))(\psi^t(y) - \psi^f(y))} \qquad (4.34)$$

and the data are assimilated in one step.

We can see not only that the number of steps in assimilating data is less, but also that the statistics required are fewer in the latter approach. Another possible advantage of using the latter approach, in which the improved estimate has the form of a Kalman filter, is that in a closed system with linear dynamics the forecast error $\psi^f - \psi^t$ forms an innovation process under moderate assumptions. So the optimal estimate field

at a certain time incorporates all the past data information. To assimilate additional data at a later time, we need only consider the new data. This allows a sequential estimation procedure. For our problem this remarkable Markov property is destroyed because the system is open and the dynamics is nonlinear. However, for simplicity, we assume this Markov property still applies. So only the data at the present time are used to form the best current estimate. The differences between these two approaches are summarized in Table 4.2.

	Our Approach	Rutherford's Approach
No. of steps of operation	2	1
Statistics required	$\overline{\psi^t}$, $\overline{\psi^x}$, $\overline{\psi^f}$ $\overline{(\psi^x-\psi^t)^2}$, $\overline{(\psi^f-\psi^t)^2}$ $\overline{(\psi^x-\psi^t)(\psi^f-\psi^t)}$ $\overline{\psi^t(x)\,\psi^t(y)}$	$\overline{\psi^t}$, $\overline{\psi^f}$ $\overline{(\psi^t(x)-\psi^f(x))(\psi^t(y)-\psi^f(y))}$
For linear and closed systems	Don't have Markov property	Have Markov property

Table 4.2 Differences between our approach and Rutherford's approach to data assimilation

Although Rutherford's approach may seem to have many advantages over our approach, we adopt ours for the following reasons: first, we have been using its first step to interpolate the observation data so

frequently. Second, we have explicitly computed only the statistics of the true field $\overline{\psi^t(x)\,\psi^t(y)}$, but not of the forecast error field $\overline{(\psi^t(x)-\psi^t(x))(\psi^t(y)-\psi^t(y))}$. (The forecast error statistics are only computed for the persistent dynamical forecast in our study.) Third, the system we are dealing with is nonlinear and open so the third advantage in Rutherford's does not survive. Here we present two cases of patch updating in the interior to see how the interior updating can improve the forecast. The first case performs a dynamical forecast by using a perfect initial condition but a persistent boundary condition; thus, the forecast error comes mainly from the boundary condition errors. The second case performs the dynamical forecast by using a zero initial condition but a perfect boundary condition; here the forecast error comes mainly from the initial condition error. The data patch we used is shown in Figure 4.7. The data (streamfunction only) are collected every 3 grid points (46.875 km) in a central square region whose size is 281.25 km square. The error model we used is shown in Figure 4.8. The 500 km square domain is divided into two regions--a patch and a rim, and the statistics in each region are treated as uniform. Figure 4.9 shows the NRMS streamfunction error as a function of time in the first case: the interior is updated every 5 time steps. We can see that the interior updating improves the estimate at the time of the data assimilation. But the forecast immediately becomes worse again after just one time step. The envelope of the error curve is essentially that of the persistence experiment. In Section II.5, where a whole field was updated, the results are similar. The suggestion is, then, that the boundary conditions have a dominant

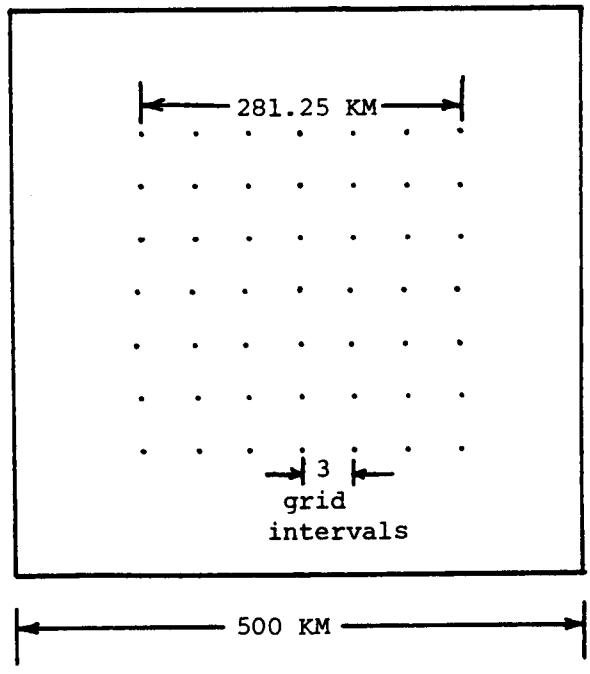


Figure 4.7 The data patch used in the interior updating.

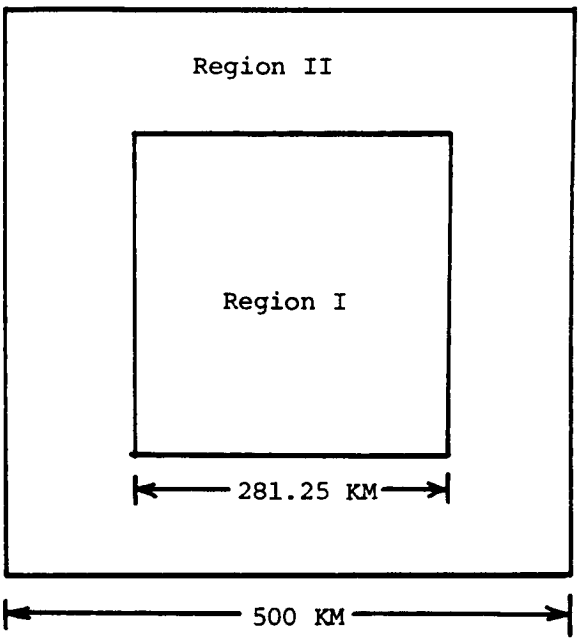


Figure 4.8 The error model used in the interior updating experiments.

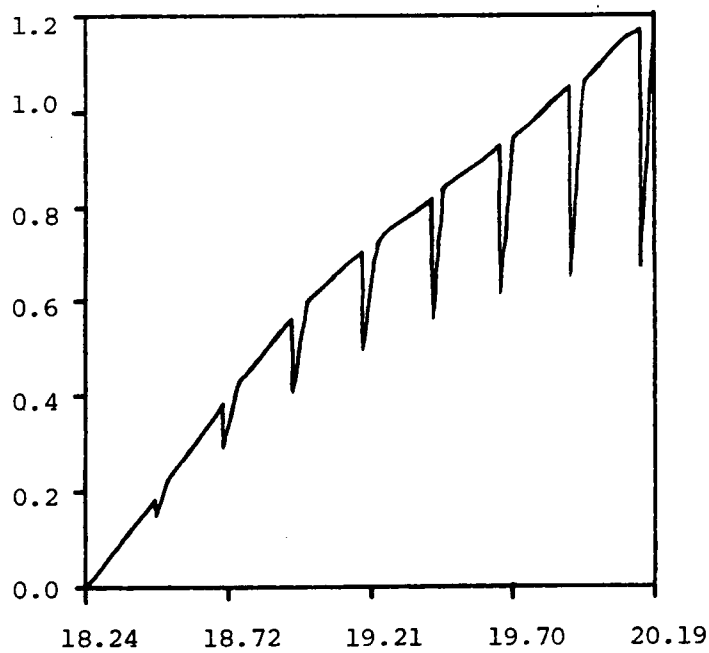


Figure 4.9 The NRMS streamfunction error in the interior updating dynamical forecast experiment in which a perfect initial condition but a persistent boundary condition is used.

influence on the dynamical forecast; the interior updating will do no good if the boundary condition itself fails to be updated. Here we confirm the same result. Figure 4.10 shows the NRMS streamfunction errors of the experiments in the second case with different updating frequencies. Curve (a) is for the experiment without updating. Curve (b), for the experiment which updates the interior every 30 time steps. Curve (c), 15 time steps, curve (d), 5 time steps. It can be seen that the updating is the most effective when it is applied for the first time. Later applications do not improve the result as much as the first one. Curves (b), (c), and (d) come very close together near the end of the experiments. Figure 4.11 shows the NRMS streamfunction errors of the experiments in the second case, however, all experiments update the interior only once. Curve (a) is for updating at time 0, curve (b) is for updating after 15 time steps and curves (c) - (e), for updating after 30, 45 and 90 time steps respectively. They all show improvements immediately from the updating, while the errors stay fairly flat afterwards. Figure 4.12 shows some typical dynamical forecast, interpolated, estimated error fields processed in the interior updating. Figure 4.12a shows the dynamical forecast error fields. Figure 4.12b shows the interpolated observation error field. Fig. 4.12c shows the optimal estimate error fields.

In all the experiments in the second case, when we update the interior streamfunction, we also update the vorticity by taking the Laplacian of the updated streamfunction using a

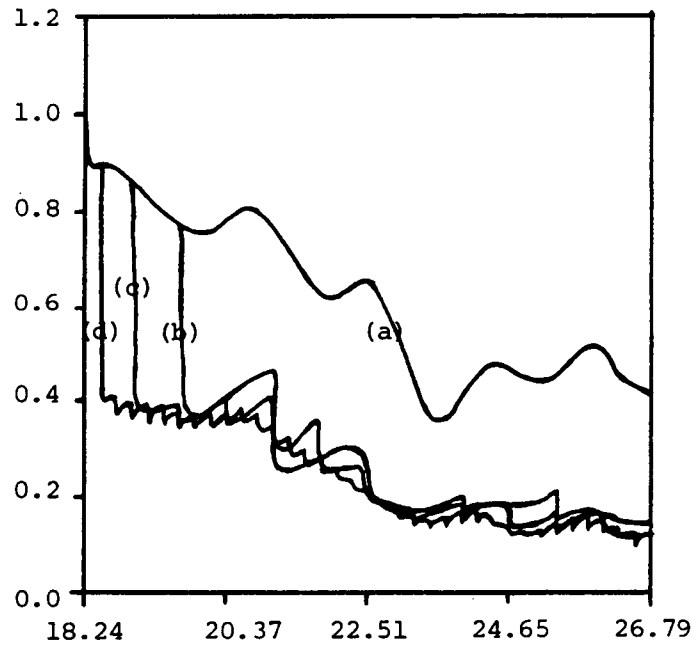


Figure 4.10 The NRMS streamfunction errors in the interior updating dynamical forecast experiments in which a zero initial condition but a perfect boundary condition is used. The interior is updated every (a) --- (b) 30 (c) 15 (d) 5 time steps.

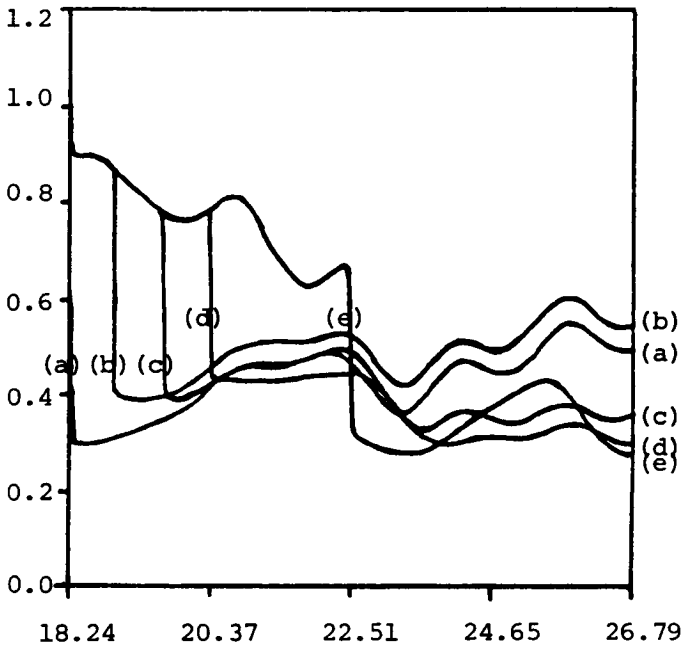
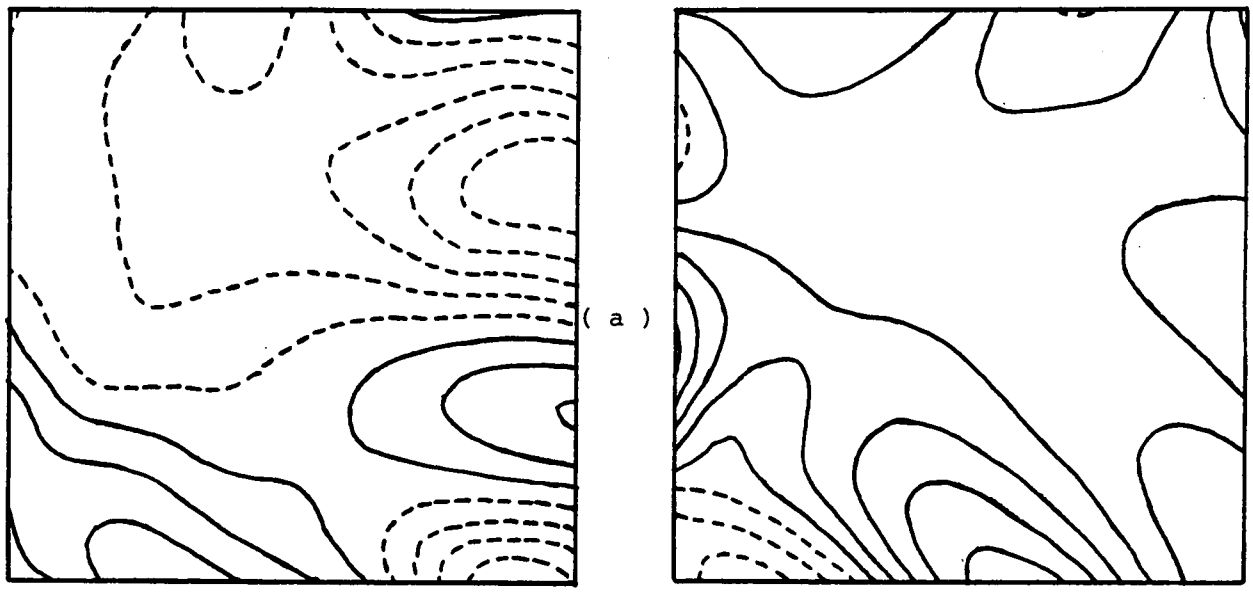


Figure 4.11 The NRMS streamfunction errors in the interior updating dynamical forecast experiments in which a zero initial condition but a perfect boundary condition is used. The interior is updated only once at
(a) 0 (b) 15 (c) 30 (d) 45 (e) 90 time steps.


(a)
Min=-2.9 Max 2.3 CI=0.6
Min=-2.6 Max=3.2 CI=0.7

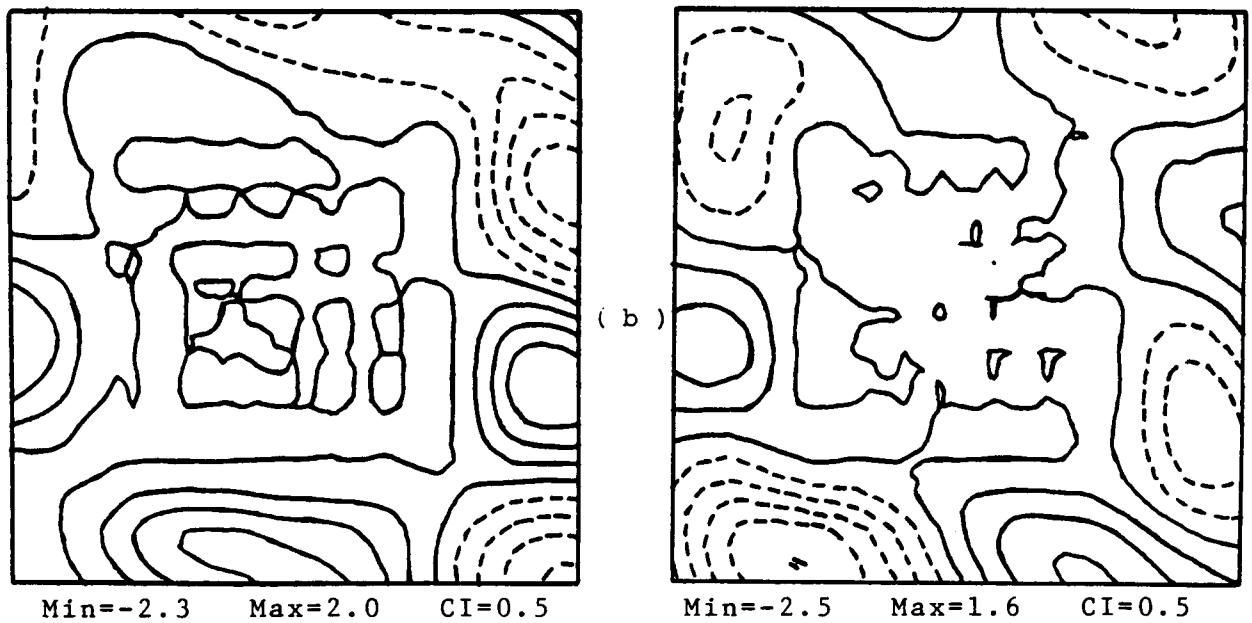

(b)
Min=-2.3 Max=2.0 CI=0.5
Min=-2.5 Max=1.6 CI=0.5

Figure 4.12

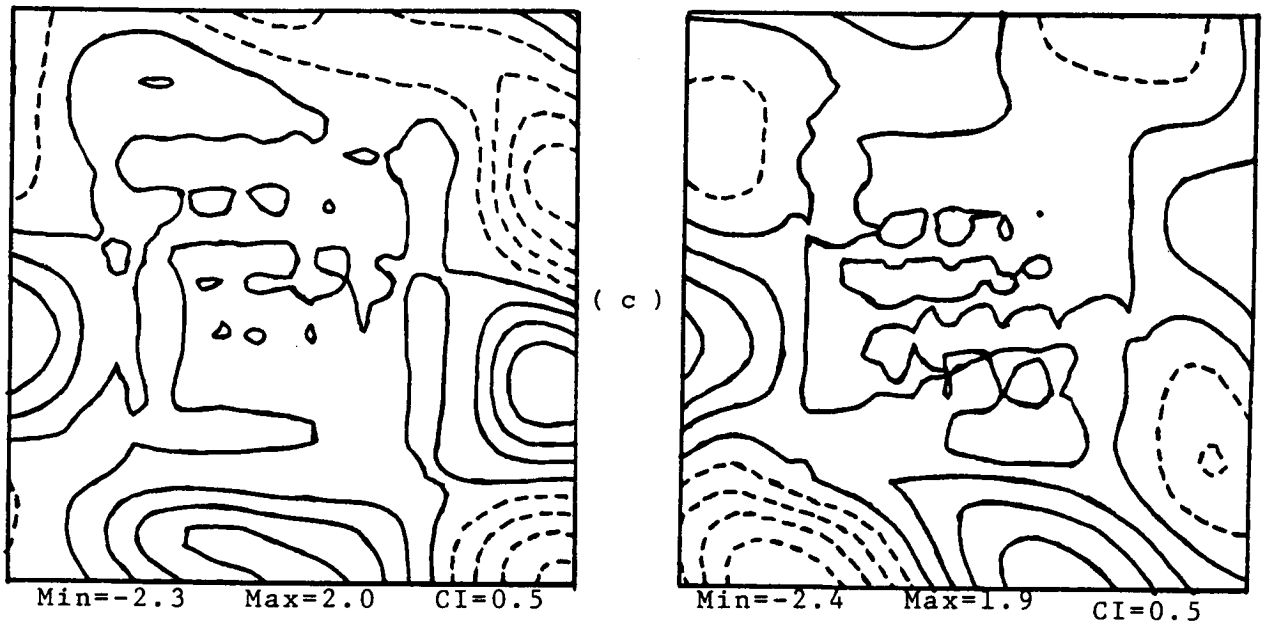


Figure 4.12 (Cont.)
Some of the error fields occur in the interior updating experiments.
(a) dynamical forecast error fields.
(b) interpolated observation error fields.
(c) optimal estimate error fields.
The figures on the left and right hand sides are the error fields of two different realizations.

2nd order scheme. If we fail to update the vorticity, then the result will not improve. As a matter of fact, it becomes worse as shown in Figure 4.13. This shows the importance of updating the vorticity in the interior in order to make the fields dynamically balanced. This point is obscured in the first case, because the result is dominated by the persistent boundary condition.

(iv) Optimal Combination of Persistent Statistical and Persistent Dynamical Forecasts

Given an initial field, we can predict the future evolution of the dynamic field by either dynamical forecast or statistical forecast. The simplest and most natural dynamical forecast is one that uses the initial boundary condition to drive the dynamical forecast model, although more complex schemes are available as discussed in example (ii). Of course, we need to have some statistics about the system. The simplest and most natural statistical forecast is one that uses the initial field as the observation data to drive the statistical forecast model. Since the data used in both cases are from one realization only, they are named the persistent dynamical forecast and the persistent statistical forecast respectively.

Both persistent forecasts are estimates of the same dynamic field. They can be combined to form a better estimate of the system by the formula for the optimal combination of estimates introduced in Section IV.1. The error model we are going to use is shown in Figure 4.14. The rationale for choosing this

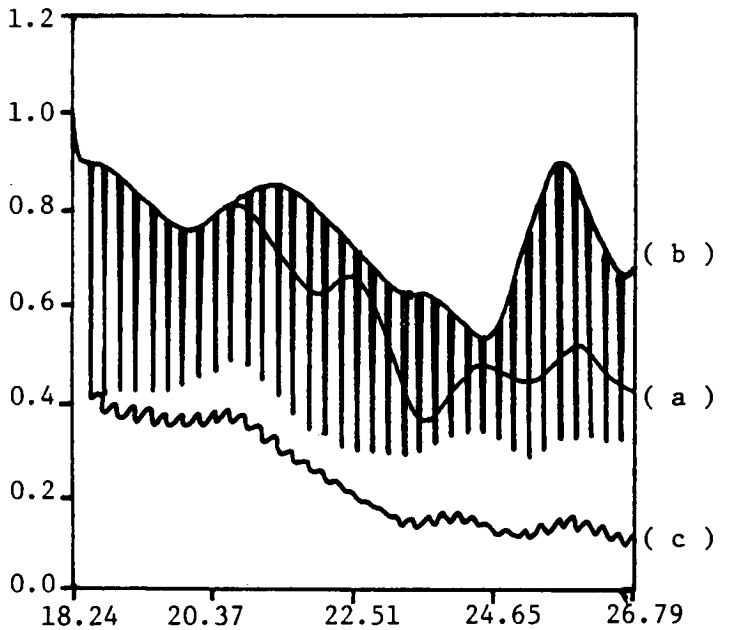


Figure 4.13 The NRMS stream function error in the interior updating dynamical forecast experiment in which a zero initial condition but a perfect boundary condition is used.
(a) No interior updating.
(b) The interior stream function is updated by a data patch every 5 time steps, but the interior vorticity is not updated.
(c) The interior stream function is updated by a data patch every 5 time steps. The interior vorticity is derived from the stream function by a second order Laplacian scheme after each updating.

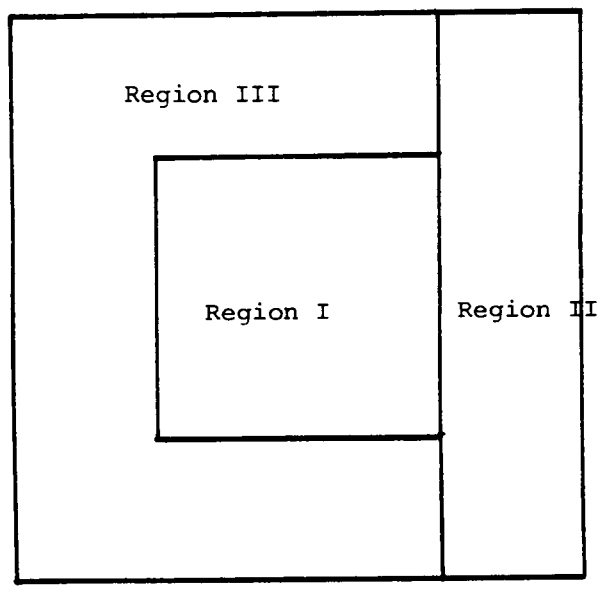


Figure 4.14 The error model used in the optimal combination of persistent statistical and persistent dynamical forecasts.

error model is that the persistent dynamical forecast has distinct error statistical properties in the central region and near the boundary as shown in Figure 2.6, while the persistent statistical forecast has distinct error statistical properties near the eastern boundary and the rest of the region as shown in Figure 3.4. This error model takes these two types of error characteristics into consideration.

In the beginning, we use the actual error statistical parameters, which are computed via the verification data. The NRMS streamfunction errors of the persistent statistical forecast, the persistent dynamical forecast, and the optimal combination of these two are shown in Figure 4.15. Curve (a), (b), (c) is for the persistent statistical forecast, the dynamical, and the optimal combination respectively. the error for the persistent statistical forecast is slightly higher than for the persistent dynamical forecast, and it is not smooth. Non-smoothness results in the statistical calculation because we have chosen the grid interval Δ h and the time step Δ t in such a way as to satisfy the Courant-Friedrichs-Lewy condition

$U \frac{\Delta t}{\Delta h} \leq \alpha$ for the dynamical calculation to be stable,

where U is the maximum of the fluid speeds and phase speeds present and α is a constant of $O(1)$ (Courant (8)). In our simulation, the simulated wave propagates westwards about 1/3 grid interval for every one time step. Thus, when it searches for the highest correlated grid point in the statistical forecast, it will sometimes have a non-smooth transition in the

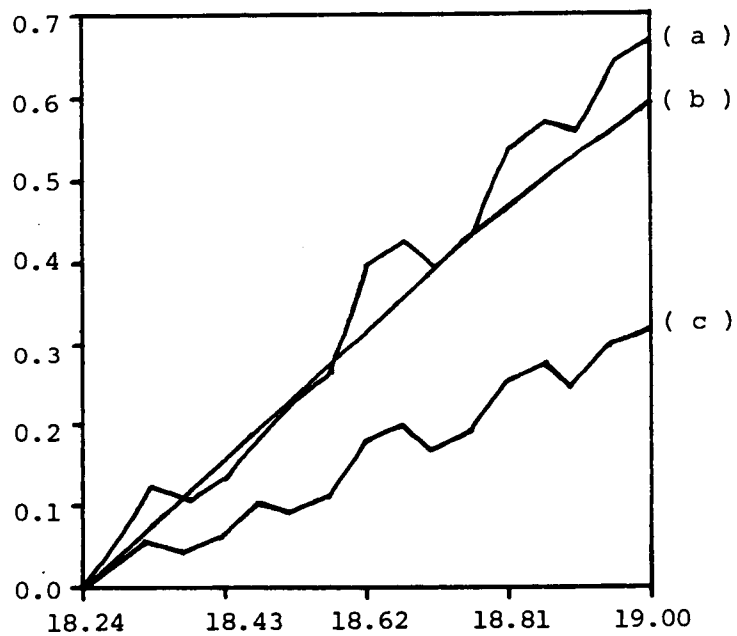


Figure 4.15 The NRMS streamfunction error of the
(a) persistent statistical forecast
(b) persistent dynamical forecast
(c) optimal combination of the persistent statistical and dynamical forecasts using actual error statistics.
The calculation starts at period 3.

highest correlated grid point selected from one time step to the next time step. And this accounts for the non-smoothness of the persistent statistical forecast error curve. It can be seen that we have contradictory objectives to be achieved here. In the dynamical calculation we want $U\frac{\Delta t}{\Delta h}$ to be less than one for the dynamical calculation to stable, while in the statistical forecast we want $U\frac{\Delta t}{\Delta h}$ to be greater than one so that the wave speeds can be resolved. The one we choose here is a compromise reached after taking both into consideration.

Next, we use the average error statistical parameters which are the averages of the error statistical parameters of several realizations. The error statistical parameters B_d (dynamical forecast bias), B_s (statistical forecast bias), E_d^2 (dynamical forecast means square error), E_s^2 (statistical forecast mean square error), γ (correlation coefficient) of 4 realizations and their average are shown in Figure 4.16 for each region in the error model and each time step. Figure 4.16a shows the error statistical parameters for the calculation that starts at period 2.25. Figure 4.16b-d, for the calculation that starts at period 3, 3.75, 4.5 respectively. Figure 4.16e is the average of Figure 4.16a-d. The NRMS streamfunction errors of the persistent statistical forecast, persistent dynamical forecast, and the optimal combination using the averaged error statistical parameters are shown again in Figure 4.17. Curves (a), (b) and (c) are for the persistent statistical forecast, persistent dynamical forecast, and the optimal combination respectively.

Figure 4.16 The error statistical parameters B_d, B_s, E_d^2, E_s^2, γ in each region at each time step for the optimal combination of persistent statistical and persistent dynamical forecast experiments that starts at period
(a) 2.25 (b) 3.00
(c) 3.75 (d) 4.50
The averages of the four realizations are shown in (e).

Time	Region	B_d	B_s	E_d^2	E_s^2	γ
	I	-0.00151	-0.05114	0.00034	0.00928	-0.59265
1	II	0.00514	0.02265	0.00604	0.01132	0.93062
	III	-0.01302	0.00741	0.00164	0.00467	0.47560
	I	-0.00001	-0.10138	0.00154	0.03860	-0.61284
2	II	0.00946	0.04628	0.02516	0.04348	0.93536
	III	-0.02290	0.01690	0.00678	0.01857	0.48978
	I	0.00318	-0.03682	0.00382	0.01750	0.30803
3	II	0.01168	-0.01182	0.05861	0.03721	-0.13565
	III	-0.03001	-0.05613	0.01580	0.01557	-0.31005
	I	0.00754	-0.08373	0.00746	0.00644	-0.18341
4	II	0.01112	0.01115	0.10765	0.03527	0.42239
	III	-0.03443	-0.04227	0.02922	0.00746	-0.04039
	I	0.01238	-0.13049	0.01272	0.01677	-0.59754
5	II	0.00712	0.03350	0.17357	0.05912	0.75198
	III	-0.03643	-0.02606	0.04756	0.00946	0.34367
	I	0.01708	-0.08150	0.01990	0.07203	0.10521
6	II	-0.00094	-0.00540	0.25758	0.17783	0.14135
	III	-0.03618	-0.09358	0.07143	0.06357	-0.18301
	I	0.02096	-0.12559	0.02931	0.04407	-0.11881
7	II	-0.01352	0.00278	0.36091	0.15606	0.30679
	III	-0.03402	-0.07344	0.10147	0.04680	-0.03860
	I	0.02325	-0.37229	0.04128	0.40790	0.14412
8	II	-0.03123	0.04835	0.48456	0.34022	0.28946
	III	-0.03030	0.04540	0.13831	0.39213	-0.12040
	I	0.02336	-0.41217	0.05616	0.37768	0.04954
9	II	-0.05418	0.06082	0.62949	0.35998	0.47291
	III	-0.02538	0.06693	0.18258	0.35215	-0.06812
	I	0.02060	-0.18864	0.07435	0.13881	-0.23059
10	II	-0.08278	-0.00393	0.79636	0.42445	0.42379
	III	-0.01963	-0.08279	0.23492	0.13620	0.04454
	I	0.01451	-0.23293	0.09619	0.12676	-0.40089
11	II	-0.11701	-0.00178	0.98560	0.46580	0.55057
	III	-0.01330	-0.05589	0.29588	0.13023	0.15704
	I	0.00455	-0.48040	0.12208	0.49021	-0.03475
12	II	-0.15679	0.01729	1.19741	0.64422	0.56037
	III	-0.00681	0.07946	0.36593	0.47933	-0.02530
	I	-0.00946	-0.51871	0.15224	0.48409	-0.12967
13	II	-0.20189	0.00705	1.43167	0.72721	0.64369
	III	-0.00031	0.10643	0.44556	0.45571	0.01667
	I	-0.02777	-0.32093	0.18764	0.28939	-0.44855
14	II	-0.25198	-0.03958	1.68791	0.88234	0.61004
	III	0.00601	-0.04058	0.53509	0.28670	0.19531
	I	-0.05088	-0.56407	0.22802	0.66083	-0.16114
15	II	-0.30676	-0.04013	1.96506	1.07580	0.63878
	III	0.01153	0.10330	0.63468	0.64685	0.04023
	I	-0.07879	-0.59993	0.27393	0.65479	-0.23405
16	II	-0.36564	-0.05314	2.26188	1.21214	0.70190
	III	0.01599	0.12934	0.74442	0.62370	0.08966

(a)

Figure 4.16 (Cont.)

Time	Region	B_d	B_s	E_d^2	E_s^2	γ
	I	0.04746	0.06928	0.00025	0.01219	-0.28736
1	II	0.00677	0.02052	0.00467	0.01233	0.67788
	III	0.04776	0.05560	0.00813	0.02185	0.81993
	I	0.10040	0.14931	0.00100	0.04906	-0.24828
2	II	0.01698	0.04609	0.01787	0.05053	0.68658
	III	0.12270	0.11413	0.03294	0.08717	0.81909
	I	0.15917	0.23078	0.00227	0.01224	0.49372
3	II	0.03039	0.25011	0.03841	0.03504	-0.07239
	III	0.16433	0.13926	0.07503	0.02383	-0.25792
	I	0.22265	0.32463	0.00411	0.01649	0.15691
4	II	0.04642	0.27184	0.06514	0.04243	0.28410
	III	0.23161	0.21018	0.13484	0.03354	0.43280
	I	0.28993	0.42664	0.00656	0.04674	-0.02849
5	II	0.06460	0.29605	0.09699	0.07649	0.48733
	III	0.30357	0.28196	0.21266	0.08863	0.66506
	I	0.36008	0.53210	0.00970	0.06099	0.54957
6	II	0.08461	0.48310	0.13299	0.10022	0.29916
	III	0.37927	0.31502	0.30866	0.10271	-0.21697
	I	0.43210	0.64163	0.01360	0.05700	0.41029
7	II	0.10597	0.54946	0.17231	0.20226	0.17060
	III	0.45766	0.39416	0.42287	0.10410	0.14374
	I	0.50510	0.99934	0.01834	0.35170	0.20749
8	II	0.12850	0.68704	0.21421	0.41914	0.13279
	III	0.53775	0.27328	0.55532	0.45105	0.11664
	I	0.57812	1.11055	0.02402	0.34670	0.14353
9	II	0.15195	0.69801	0.25812	0.47507	0.22373
	III	0.61850	0.35473	0.70584	0.50098	0.26389
	I	0.65045	0.99120	0.03074	0.15016	0.46670
10	II	0.17627	0.76378	0.30361	0.42311	0.29463
	III	0.69907	0.58968	0.87442	0.25017	0.08139
	I	0.72128	1.10803	0.03861	0.16779	0.37433
11	II	0.20141	0.78791	0.35036	0.57116	0.29857
	III	0.77850	0.66565	1.06083	0.30107	0.28058
	I	0.78983	1.43461	0.04775	0.44083	0.26622
12	II	0.22728	0.88604	0.39820	0.86628	0.28903
	III	0.85583	0.56140	1.26485	0.72436	0.22890
	I	0.85589	1.54308	0.05827	0.46348	0.22572
13	II	0.25413	0.90512	0.44708	1.04962	0.27793
	III	0.93072	0.63536	1.48636	0.83393	0.32588
	I	0.91883	1.45379	0.07030	0.31654	0.42435
14	II	0.28211	0.95855	0.49708	1.13909	0.28877
	III	1.00229	0.84667	1.72515	0.54533	0.24880
	I	0.97837	1.73970	0.08398	0.59050	0.35214
15	II	0.31134	1.02023	0.54839	1.48715	0.29022
	III	1.07012	0.75578	1.98104	1.05300	0.23621
	I	1.03424	1.83697	0.09945	0.61541	0.32567
16	II	0.34203	1.01470	0.60126	1.71721	0.29437
	III	1.13369	0.81915	2.25391	1.16630	0.32046

(b)

Figure 4.16 (Cont.)

Time	Region	B_d	B_s	E_d^2	E_s^2	γ
	I	0.07853	0.13982	0.00205	0.02891	-0.69487
1	II	0.16204	0.09960	0.00603	0.01141	0.71540
	III	0.04429	0.03163	0.00846	0.01501	0.79004
	I	0.16817	0.27466	0.00856	0.11406	-0.68257
2	II	0.33047	0.19514	0.02298	0.04541	0.68495
	III	0.08645	0.05969	0.03410	0.06148	0.78811
	I	0.26360	0.26102	0.01973	0.04635	0.85425
3	II	0.50011	0.20506	0.04921	0.04856	0.19033
	III	0.12500	0.27478	0.07726	0.02183	-0.22120
	I	0.36442	0.37388	0.03558	0.01809	0.39911
4	II	0.66916	0.30157	0.08300	0.06011	0.43031
	III	0.15992	0.29155	0.13806	0.02308	0.41799
	I	0.46993	0.47882	0.05589	0.05338	-0.30887
5	II	0.83550	0.39026	0.12265	0.09233	0.53223
	III	0.19098	0.30222	0.21642	0.05730	0.67797
	I	0.57943	0.48295	0.08022	0.20447	0.79358
6	II	0.99732	0.36584	0.16653	0.18802	0.31899
	III	0.21807	0.46686	0.31205	0.08066	-0.10837
	I	0.69180	0.55821	0.10792	0.12618	0.61024
7	II	1.15255	0.41396	0.21326	0.22834	0.24782
	III	0.24072	0.46284	0.42438	0.07964	0.22336
	I	0.80597	0.53412	0.13819	0.42528	0.48885
8	II	1.29945	0.36920	0.26160	0.60276	0.30113
	III	0.25861	0.57029	0.55268	0.57826	0.17697
	I	0.92084	0.58523	0.17008	0.36270	0.28614
9	II	1.43638	0.43594	0.31062	0.62515	0.32413
	III	0.27144	0.54612	0.69606	0.64054	0.28321
	I	1.03511	0.66331	0.20262	0.36779	0.53094
10	II	1.56181	0.47072	0.35973	0.50594	0.22101
	III	0.27884	0.51208	0.85339	0.20913	0.22549
	I	1.14763	0.67643	0.23477	0.33654	0.30322
11	II	1.67444	0.50684	0.40866	0.55546	0.18467
	III	0.28051	0.47071	1.02350	0.26071	0.38230
	I	1.25715	0.62787	0.26558	0.64201	0.29856
12	II	1.77317	0.45814	0.45750	0.93362	0.20856
	III	0.27616	0.49575	1.20509	0.87000	0.26428
	I	1.36264	0.61009	0.29416	0.64959	0.11490
13	II	1.85711	0.47570	0.50683	0.95187	0.16154
	III	0.26567	0.42538	1.39689	0.98645	0.32601
	I	1.46296	0.61423	0.31971	0.70028	0.23613
14	II	1.92549	0.44678	0.55748	0.92726	0.04893
	III	0.24893	0.34788	1.59762	0.54558	0.36117
	I	1.55709	0.54251	0.34170	1.04670	0.22661
15	II	1.97783	0.39772	0.61055	1.26288	0.06067
	III	0.22580	0.30473	1.80606	1.24771	0.27168
	I	1.64448	0.46596	0.35971	1.06315	0.08870
16	II	2.01399	0.39668	0.66761	1.26731	0.00847
	III	0.19675	0.21012	2.02129	1.38664	0.32618

(c)

Figure 4.16 (cont.)

Time	Region	B_d	B_s	E_d^2	E_s^2	γ
	I	-0.12245	-0.22445	0.00164	0.02282	-0.00195
1	II	-0.10613	-0.00487	0.00510	0.01173	0.69212
	III	-0.12305	-0.07700	0.01068	0.02693	0.66452
	I	-0.24361	-0.44660	0.00627	0.09145	0.03380
2	II	-0.22384	-0.01208	0.02048	0.04746	0.69021
	III	-0.23663	-0.15244	0.04240	0.10814	0.68888
	I	-0.36637	-0.43273	0.01357	0.04727	0.85366
3	II	-0.35352	-0.43584	0.04610	0.02561	0.50978
	III	-0.34254	-0.41499	0.09422	0.08955	0.40844
	I	-0.49157	-0.65662	0.02328	0.06836	0.74255
4	II	-0.49485	-0.44944	0.08187	0.03979	0.77352
	III	-0.44054	-0.47744	0.16496	0.11311	0.69446
	I	-0.61985	-0.87515	0.03524	0.13676	0.55399
5	II	-0.64733	-0.46305	0.12767	0.07992	0.79681
	III	-0.53020	-0.53557	0.25342	0.19053	0.80356
	I	-0.75283	-0.86535	0.04931	0.20115	0.84748
6	II	-0.81074	-0.80956	0.18339	0.08449	0.63110
	III	-0.61235	-0.80737	0.35839	0.33516	0.40956
	I	-0.89145	-1.07763	0.06538	0.21547	0.81485
7	II	-0.98430	-0.82864	0.24879	0.11057	0.62058
	III	-0.68748	-0.84902	0.47862	0.35426	0.57695
	I	-1.03649	-1.21274	0.08330	0.48490	0.66245
8	II	-1.16710	-0.85019	0.32376	0.18684	0.44440
	III	-0.75599	-0.94323	0.61301	0.61219	0.38650
	I	-1.18852	-1.41000	0.10290	0.58850	0.58870
9	II	-1.35807	-0.86667	0.40792	0.24170	0.51118
	III	-0.81828	-0.97148	0.76052	0.66385	0.50636
	I	-1.34794	-1.47951	0.12397	0.45492	0.79913
10	II	-1.55607	-1.11865	0.50102	0.20924	0.59654
	III	-0.87476	-1.15253	0.92026	0.70222	0.51369
	I	-1.51497	-1.65797	0.14615	0.50733	0.70144
11	II	-1.75942	-1.13149	0.60253	0.27065	0.58811
	III	-0.92603	-1.16301	1.09150	0.76182	0.60621
	I	-1.68944	-1.77219	0.16903	0.84468	0.58940
12	II	-1.96651	-1.14726	0.71206	0.35216	0.46791
	III	-0.97238	-1.20697	1.27360	0.99895	0.48751
	I	-1.87135	-1.92559	0.19218	0.98797	0.48916
13	II	-2.17595	-1.15598	0.82917	0.44148	0.44831
	III	-1.01449	-1.20385	1.46597	1.09473	0.53450
	I	-2.06018	-0.97942	0.21496	0.78542	0.63964
14	II	-2.38554	-1.31104	0.95303	0.44462	0.47174
	III	-1.05289	-1.33869	1.66816	1.14991	0.53638
	I	-2.25559	-2.06369	0.23680	1.13802	0.52545
15	II	-2.59389	-1.31559	1.08319	0.53045	0.38625
	III	-1.08830	-1.34123	1.87971	1.37728	0.45451
	I	-2.45626	-2.17075	0.25715	1.28110	0.41367
16	II	-2.79870	-1.31158	1.21894	0.63520	0.37132
	III	-1.12061	-1.30900	2.10021	1.43997	0.51167

(d)

Figure 4.16 (cont.)

Time	Region	B_d	B_s	E_d^2	E_s^2	γ
	I	0.00051	-0.01662	0.00107	0.01830	-0.39421
1	II	0.01696	0.03450	0.00546	0.01170	0.75401
	III	-0.01100	0.00441	0.00723	0.01712	0.68775
	I	0.00624	-0.03100	0.00434	0.07329	-0.37747
2	II	0.03327	0.06886	0.02162	0.04672	0.74928
	III	-0.01759	0.00957	0.02906	0.06884	0.69647
	I	0.01489	0.00556	0.00985	0.03084	0.62742
3	II	0.04716	0.00188	0.04808	0.03661	0.12302
	III	-0.02080	-0.01427	0.06558	0.03770	-0.09518
	I	0.02576	-0.01046	0.01761	0.02735	0.27879
4	II	0.05796	0.03378	0.08441	0.04440	0.47758
	III	-0.02086	-0.00449	0.11677	0.04430	0.37622
	I	0.03810	-0.02505	0.02760	0.06341	-0.09523
5	II	0.06497	0.06419	0.13022	0.07697	0.64209
	III	-0.01802	0.00564	0.18251	0.08648	0.62257
	I	0.05094	0.01705	0.03978	0.13466	0.57396
6	II	0.06756	0.00850	0.18512	0.13764	0.34765
	III	-0.02180	-0.02977	0.26263	0.14552	-0.02470
	I	0.06335	-0.00084	0.05405	0.11068	0.42914
7	II	0.06517	0.03479	0.24882	0.17431	0.33649
	III	-0.00578	-0.01636	0.35684	0.14620	0.22636
	I	0.07446	-0.01289	0.07028	0.41745	0.37573
8	II	0.05741	0.06360	0.32103	0.38724	0.29195
	III	0.00252	-0.01356	0.46483	0.50841	0.13993
	I	0.08345	-0.03160	0.08829	0.41890	0.26698
9	II	0.04402	0.08202	0.40514	0.42548	0.38299
	III	0.01157	-0.00093	0.58625	0.53938	0.24633
	I	0.08956	-0.00341	0.10792	0.27792	0.39154
10	II	0.02481	0.02798	0.49018	0.39069	0.38399
	III	0.02088	-0.03339	0.72075	0.32443	0.21628
	I	0.09211	-0.02661	0.12893	0.28461	0.24453
11	II	-0.00014	0.04037	0.58679	0.46577	0.40548
	III	0.02992	-0.02064	0.86793	0.36346	0.35653
	I	0.09052	-0.04753	0.15111	0.60443	0.27986
12	II	-0.03071	0.05355	0.69129	0.69907	0.37944
	III	0.03820	-0.01759	1.02737	0.76816	0.23885
	I	0.08443	-0.07278	0.17426	0.64628	0.17503
13	II	-0.06665	0.05797	0.80369	0.79255	0.38287
	III	0.04540	-0.00917	1.19870	0.84271	0.30077
	I	0.07346	-0.05808	0.19815	0.84271	0.21289
14	II	-0.10748	0.01368	0.92388	0.84833	0.35487
	III	0.05109	-0.04618	1.38151	0.63188	0.33542
	I	0.05725	-0.08639	0.22263	0.85901	0.23577
15	II	-0.15287	0.01556	1.05180	1.08907	0.34398
	III	0.05479	-0.04436	1.57537	1.08121	0.25066
	I	0.03592	-0.11694	0.24756	0.90361	0.14850
16	II	-0.20208	0.01167	1.18742	1.20797	0.34402
	III	0.05646	-0.03760	1.77996	1.15415	0.31199

(e)

Figure 4.16 (Cont.)

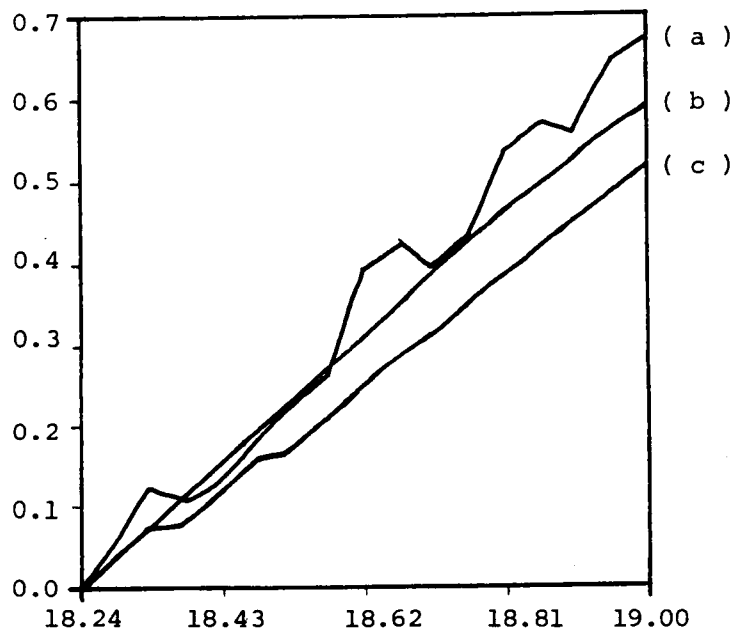


Figure 4.17 The NRMS streamfunction error of the
(a) persistent statistical forecast
(b) persistent dynamical forecast
(c) optimal combination of the persistent statistical and dynamical forecasts using averaged error statistics.
The calculation starts at period 3.

It can be seen that the NRMS error of the optimal combination using averaged error statistics is larger than that using actual error statistics. But it still shows improvement over errors in the persistent statistical and dynamical forecasts. Figure 4.18a-d shows the true field, persistent dynamical forecast field, persistent statistical forecast field, and optimal estimated field respectively after 8 time steps. Figure 4.19a-d shows the same things except after 16 time steps.

IV.4 Optimal Exploitation of One Data Realization

In this section, we are interested in answering the question: what is the best we can forecast the future, given a perfect initial field? In a closed system, the answer to this question is easy because the initial condition is the only driving force. The best forecast is simply the run of the dynamical model using this initial condition. In an open system, however, the initial condition is not the only driving force. Another driving force, the boundary condition, is unknown. In Example (ii) of Section IV.3, we knew the persistent dynamical forecast was by no means the best. We got a better result when we used a better estimate of the boundary condition. In that example, we used the boundary condition which was statistically forecasted from the initial condition and the result improved. But this is still not the best forecast. In the following paragraph, we devise some better schemes to estimate the boundary condition.

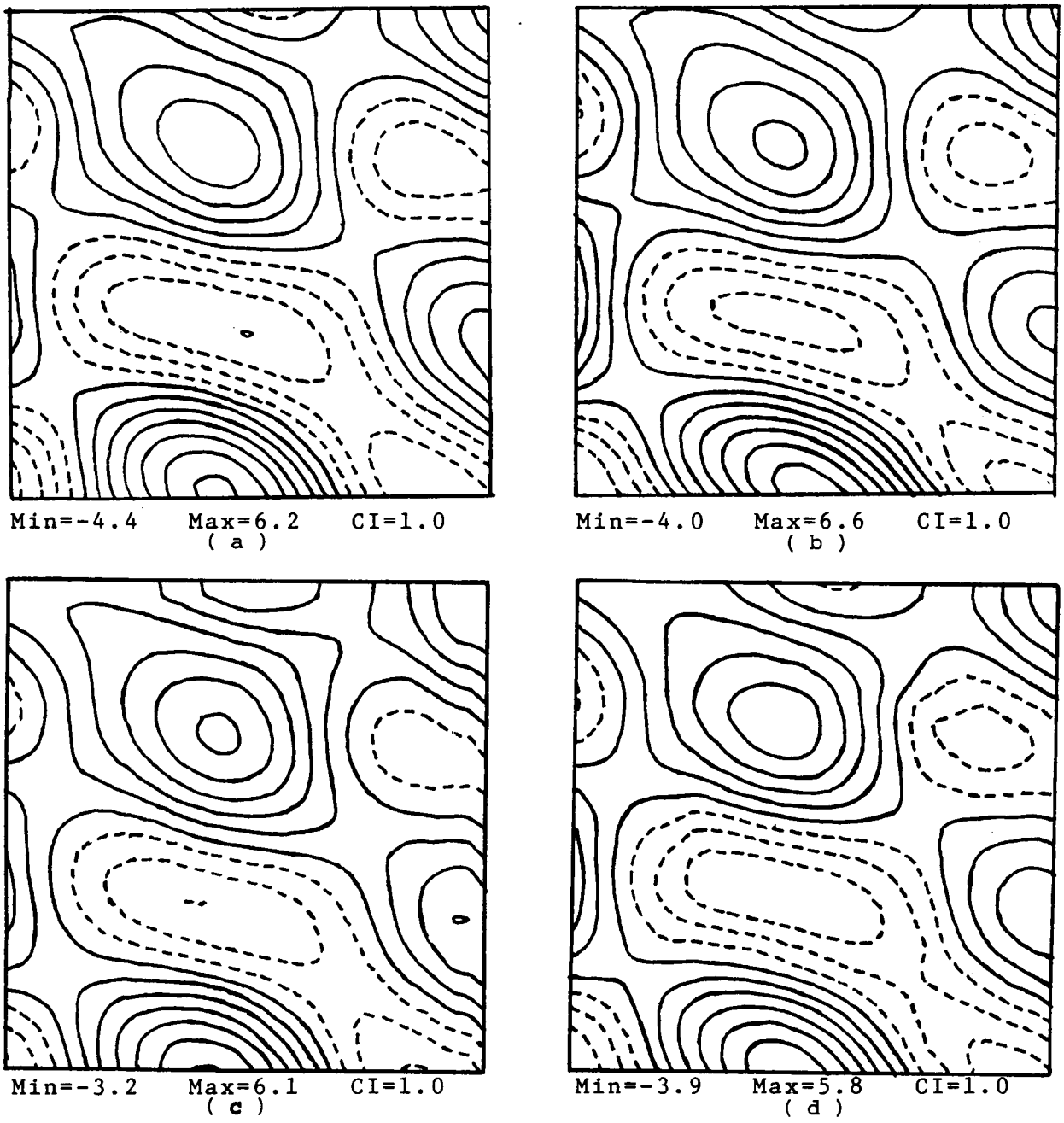


Figure 4.18 Some of the fields that are generated in the optimal combination of the persistent statistical and dynamical forecasts experiment.
(a) The true (b) The dynamical forecast
(c) The statistical forecast
(d) The optimal estimated field after 8 time steps.
The calculation starts at period 3.

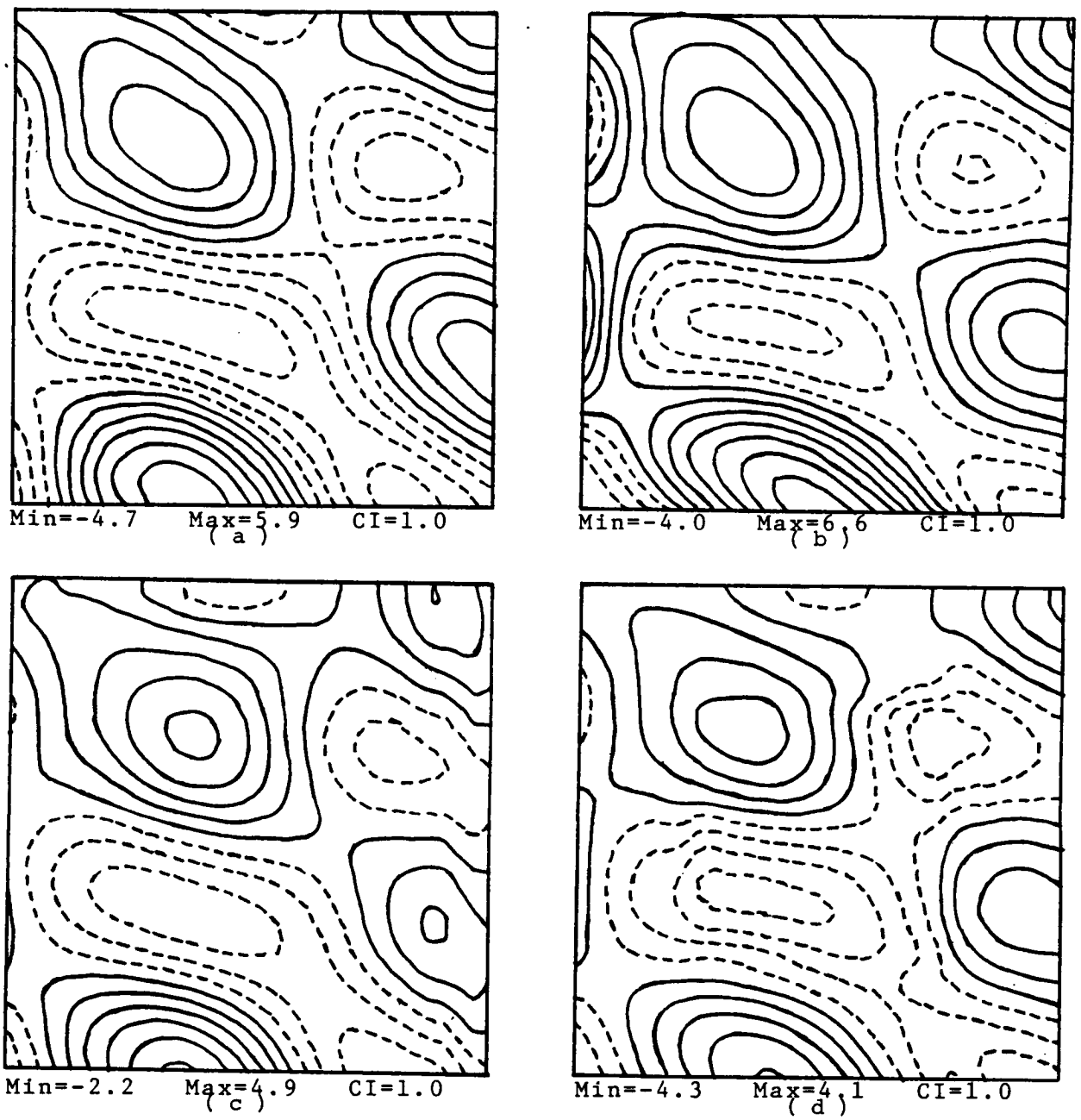


Figure 4.19 Some of the fields that are generated in the optimal combination of the persistent statistical and dynamical forecasts experiment.
(a) The true (b) The dynamical forecast
(c) The statistical forecast
(d) The optimal estimated field after 16 time steps.
The calculation starts at period 3.

As we have seen in example (ii) of Section IV.3, the statistical forecast boundary condition from the dynamical field at the previous time step is not good because of the way the grid interval and the time step are chosen. The eddy fields propagate westward about 1/3 grid interval per time step. This speed cannot be resolved by our statistical model. In the following experiment, we try a way to advect the dynamical field 1/3 grid interval to obtain the boundary condition at the next time step. This is implemented by a simple linear interpolation formula $\psi^{t+1}(x,y) = \frac{1}{3}\psi^{t}(x,y) + \frac{2}{3}\psi^{t}(x+\Delta x, y)$. The NRMS streamfunction error of the dynamical forecast using this advected boundary condition is shown in Figure 4.20(d), which is compared to those of the persistent dynamical forecast (a), dynamical forecast using the statistical forecast boundary condition from the perfect initial field (b), dynamical forecast using the statistical forecast boundary condition from the previous dynamical field (c). Although it is still worse than the persistent dynamical forecast (a), it is better than the dynamical forecast using the statistical forecast boundary condition from the previous dynamical field (c). If we use a better interpolation scheme and/or treat the wave speed more correctly, we should obtain a result better than that of the persistent dynamical forecast.

We have seen several estimates of the boundary conditions--persistence, statistical forecast from the perfect initial field, statistical forecast from the previous dynamical field, advection--which are available at the next time step. If we

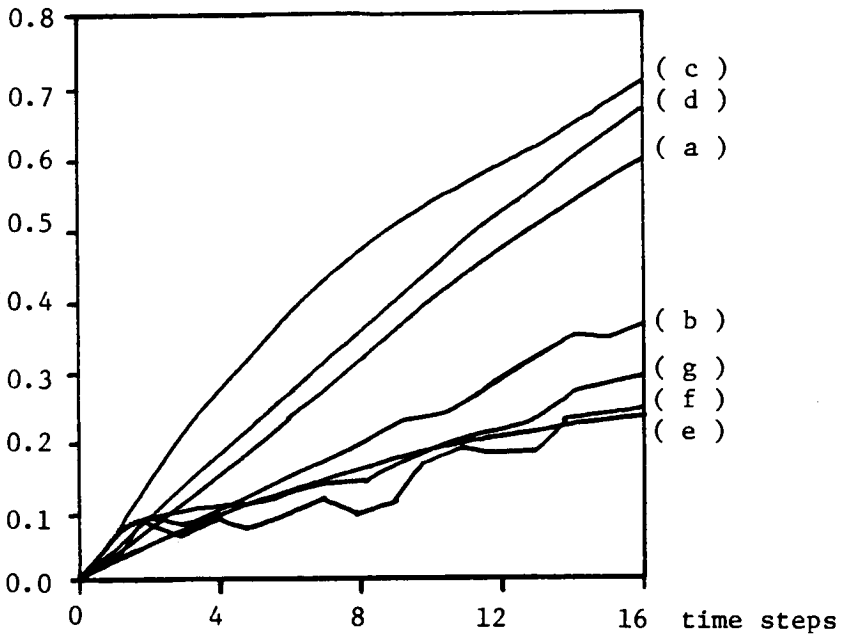


Figure 4.20 The NRMS stream function errors of the dynamical forecasts using the
(a) persistent boundary condition.
(b) statistical forecast boundary condition from the perfect initial field.
(c) statistical forecast boundary condition from the previous dynamical field.
(d) advected boundary condition from the previous dynamical field by 1/3 grid point.
(e) optimal combination of the persisted and the statistically forecasted boundary conditions from the previous dynamical field.
(f) optimal combination of the statistically forecasted boundary conditions from the perfect initial field and the previous dynamical field.
(g) optimal combination of the advected boundary condition from the previous dynamical field and the statistically forecasted boundary condition from the perfect initial field.

combine them optimally by the formulas introduced in Section IV.1, we can obtain a better estimate of the boundary condition. Figure 4.20 (e), (f), and (g) show the NRMS streamfunction errors for the dynamical forecasts using the optimal combination of the persisted and the statistically forecasted boundary conditions from the previous dynamical field, the optimal combination of the statistically forecasted boundary conditions from the perfect initial field and the previous dynamical field, the optimal combination of the advected boundary condition from the previous dynamical field and the statistically forecasted boundary condition from the perfect initial field respectively. The logic diagrams for (e), (f) and (g) are shown in Figure 4.21. The error model used for the optimal combination is shown in Figure 4.22. They are all better than any of the forecast schemes described above and are very close together. This may represent the upper limit we can do for the particular error model chosen for the optimal combination. To improve the results, we have to refine our error model to get even better estimates of the boundary condition. Of course, we have to pay additional cost--i.e., the number of the statistical parameters required will be increased.

In the extreme case, where each grid point on the boundary forms a separate region, we can get the best estimate of the boundary condition, hence the best forecast. But the number of statistical parameters required will be intolerable.

Table 4.3 summarizes all the experiments that are performed in regard to the optimal exploitation of one data realization.

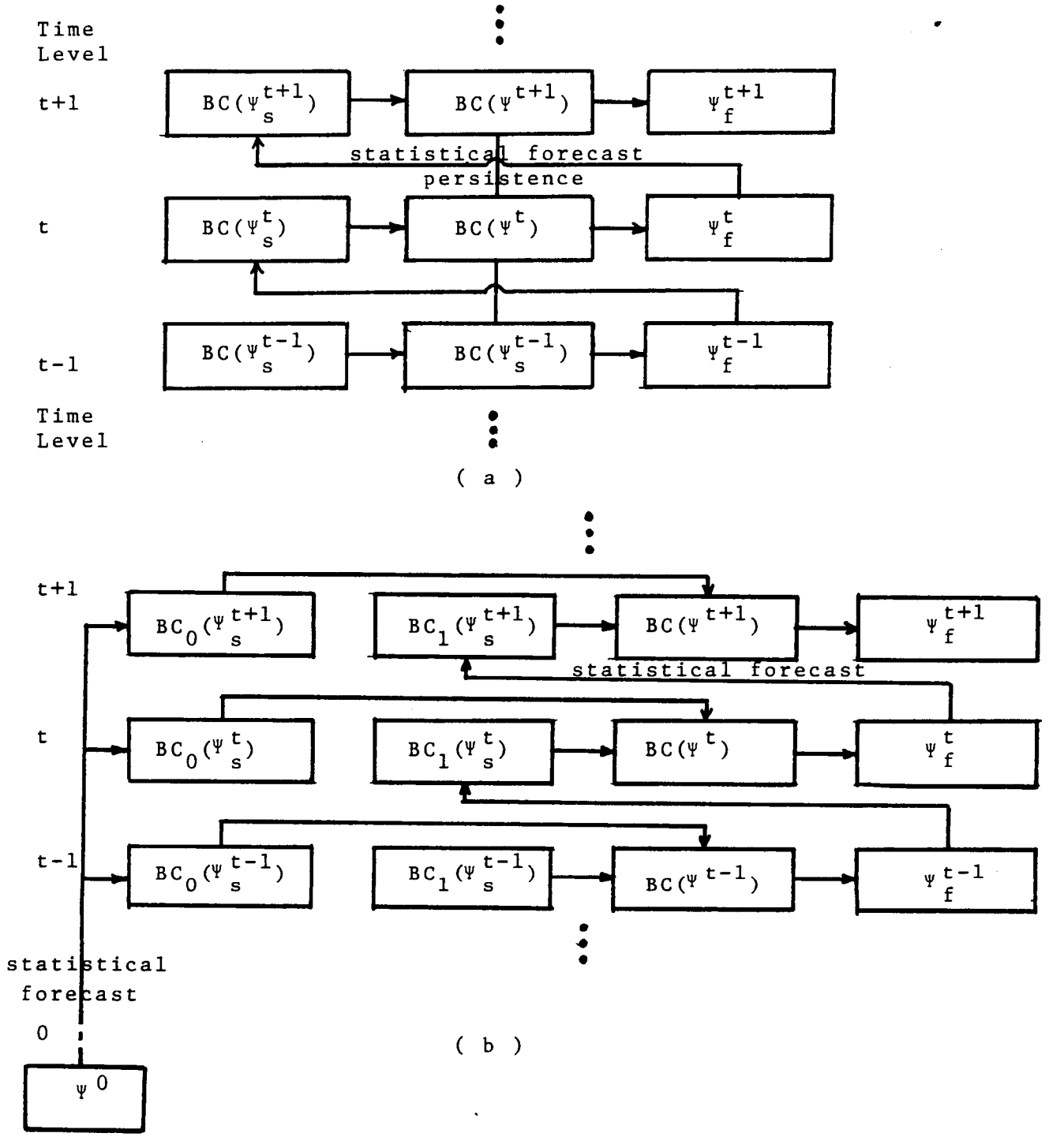

Time Level
t+1
t
t-1
Time Level
statistical forecast
persistence
(a)
statistical forecast
statistical forecast
0
(b)

Figure 4.21

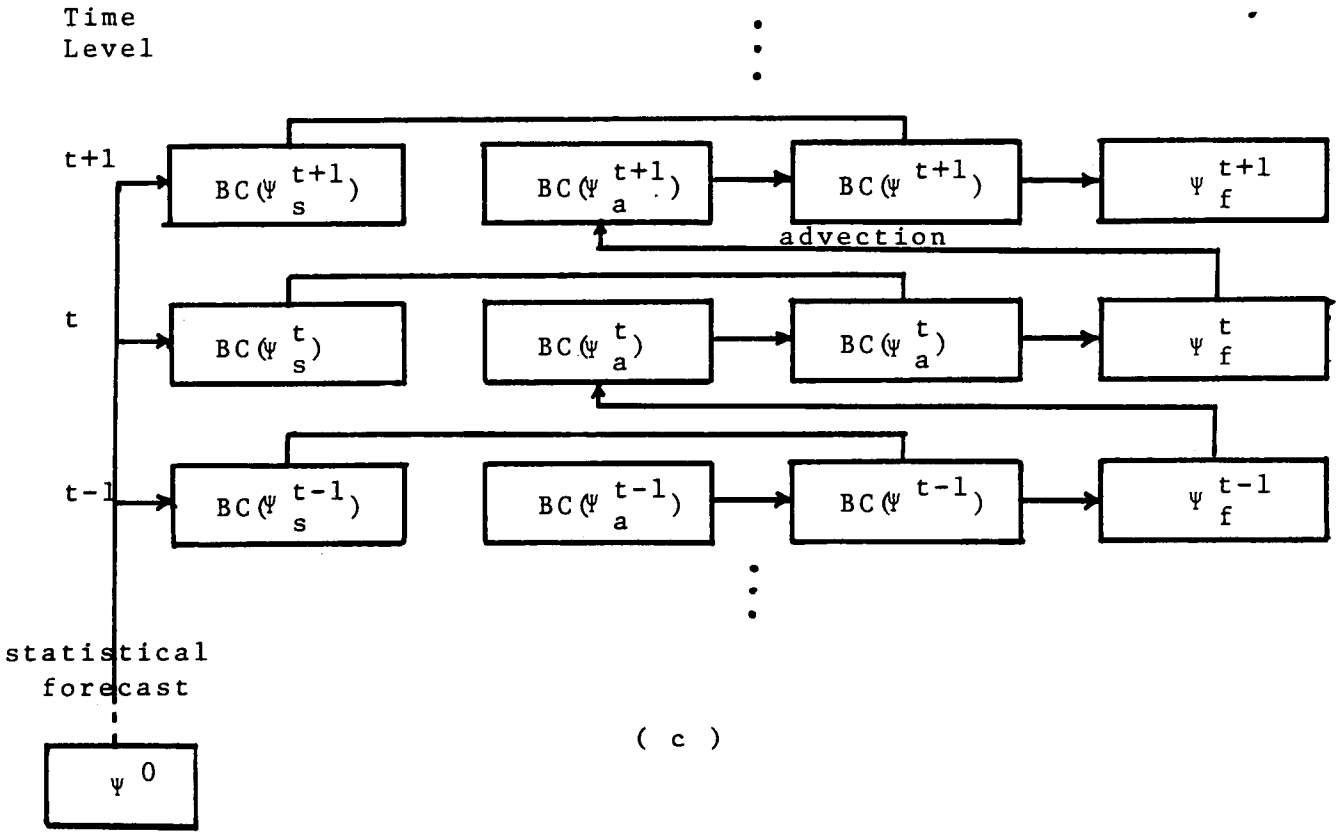


Figure 4.21 (Cont.)
The logic diagrams for experiments
(a) 4.4.5
(b) 4.4.6
(c) 4.4.7

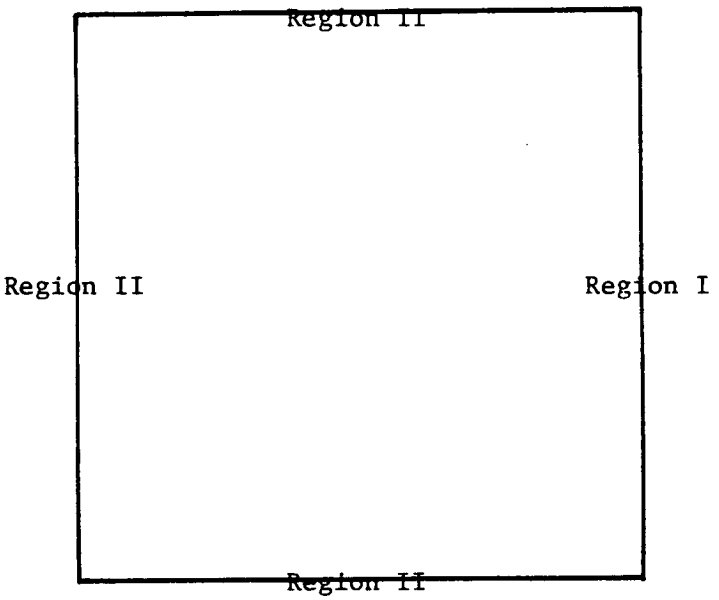


Figure 4.22 The error model used for the optimal combination of boundary conditions in experiments 4.4.5-7. The whole boundary is divided into two regions--one is the eastern boundary, the other is the rest of the boundary.

Expt. No.	Boundary condition used	*Parameters used in the Statistical Forecast	Curve No. in Fig. 4.20
4.4.1	persistence from the initial boundary condition		(a)
4.4.2	statistical forecast from the perfect initial field	$C_{max} = 0.95$ $N = 4$	(b)
4.4.3	statistical forecast from the previous dynamical field	$C_{max} = 0.95$ $N = 2$	(c)
4.4.4	advection from the previous dynamical field by 1/3 grid point		(d)
4.4.5	optimal combination of the persisted and the statistically forecasted boundary conditions from the previous dynamical field	$C_{max} = 0.95$ $N = 2$	(e)
4.4.6	optimal combination of the statistically forecasted boundary conditions from the perfect initial field and the previous dynamical field	$C_{max} = 0.95$ $N = 4$ (from the perfect initial field) $N = 2$ (from the previous dynamical field)	(f)
4.4.7	optimal combination of the advected boundary condition from the previous dynamical field and the statistically forecast boundary condition from the perfect initial field	$C_{max} = 0.95$ $N = 4$	(g)

Table 4.3 Summary of the optimal exploitation of one data realization experiments.

* The parameters C_{max} and N are the maximum correlation any two chosen data points can have and the number of data points chosen for statistical forecast, respectively as defined in Section III.3.

Of course, we could have applied the optimal combination to the interior to improve the forecast. However, we learned from previous results that the interior updating can only improve the estimate at the time of updating. It does not help in any forecasting which is essentially dominated by the boundary condition. Also, it is very costly, so we decided in the end not to do it here.

CHAPTER V
CONCLUSIONS AND FUTURE RESEARCH SUGGESTIONS

In this thesis, we have tried various approaches to forecast the evolution of the mesoscale currents in a block of ocean. The dynamical approach bases the forecast on solving the initial and boundary value problem posed by the linearly damped barotropic potential vorticity equation. A finite element method using rectangular bilinear elements is used to solve the above equation numerically. Because it is capable of delivering remarkably accurate forecast results if provided with perfect initial and boundary data, it is adopted as the "Harvard Open Ocean Dynamical Model." Boundary conditions for the dynamical model are specially treated because of the open boundary on which the boundary conditions vary in time with the flow. Some sufficient conditions are proposed for a set of open boundary conditions to be well-posed, but what the necessary and sufficient conditions are for the well-posedness is unknown. I suggest the research along this line be continued in the future. Close collaborations with mathematicians are expected.

We obtain an estimate of the error of the solution in terms of the errors of the initial condition and the boundary conditions accumulated in time as a by-product in studying the open boundary condition problem. But the estimate is usually too conservative. The bound in the estimate is precise only when the dynamics are linear and the boundary conditions are per-

fect (i.e., error free).

Some benchmark calculations were performed to provide comparison data for the later forecast experiments. The most interesting benchmark calculations are the calculation that uses perfect initial and boundary data and the calculation that uses perfect initial data and persistent boundary data.

To improve over persistent dynamical forecast, we can update the boundary condition and/or the interior field. (There is a difference between updating the interior and reinitialization. Reinitialization not only updates the interior, but also updates the boundary condition.) It was found that the NRMS streamfunction error was essentially dominated by the boundary condition updating. The interior updating did not do much good if the boundary condition failed to be updated. The effect due to interior updating can only be seen after some time has passed (about a month). This suggested that the most efficient way of collecting data to do a good forecast was to collect the streamfunction data on the boundary which was further substantiated by later experiments. (The forecast is relatively insensitive to the boundary vorticity data -- Robinson and Haidvogel, (26).

The statistical approach bases the forecast on the space-time objective analysis formula. The correlation function used was computed from the simulated ocean data set generated by the exterior dynamical calculation. In the future, this should be computed from the real ocean data set. A data point selecting algorithm was devised to pick the obervation data for the in-

terpolation to be stable and efficient. Many statistical forecast experiments which use data at one time level were performed. These results were used to find the best parameters to be used in the statistical model. A statistical forecast experiment which uses data at different time levels is also shown. Here we devise a simple cruising path and data sampling scheme for a ship steaming in the mid-ocean. It is found that the error is maintained at about 30 percent after one lap of the curising. What the optimal cruising path and data sampling scheme is for the field to be maintained at a certain level of accuracy is suggested for future studies.

The combined statistical and dynamical approach bases the forecast on the combination of the above two approaches. The statistical model is used to provide the initial and boundary conditions and to assimilate recently available data into the dynamical model. It would be interesting to know what the most efficient way of collecting the initial and boundary data is in order that the forecast satisfies a given accuracy requirement. Example IV.3.1 gives some clues to the answer. But there is still a lot of work to be done before the results can be really useful.

Some interior patch updating experiments were also performed in which a patch of new data is assimilated into the dynamical model. The reinitialization experiments above were special cases of this in which the whole field is updated. It is found that the patch updating can be used to improve the

dynamical forecast due to the initial condition error but not the boundary condition error.

Next we turned our attention to what the best forecast is, given a perfect initial field. Since the boundary condition has a significant influence on the forecast result, it is important to obtain a good estimate of the boundary condition. First we tried using the boundary conditions statistically forecasted from the perfect initial field and the previous dynamical field. Both showed worse than the persistent dynamical forecast in the first few time steps. This is because the grid interval Δx and the time step at Δt chosen are not able to resolve the eddy wave speeds. The dynamical forecast using the statistical forecast boundary condition from the perfect initial field improves afterwards, but from the previous dynamical field is always worse than the persistent dynamical forecast. To remedy this problem, next we tried advecting the dynamical field 1/3 grid point to obtain the boundary condition at the next time step. The result improved, but it is still worse than the persistence. We suggest that a better interpolation scheme and/or a more accurate wave speed value be used in the future to do advection. The result should be able to be made better than the persistence. Finally, we tried three different ways of estimating the boundary conditions from two a priori estimates. The first is to combine the persisted and the statistically forecasted boundary conditons from the previous dynamical field. The second is to combine the statistically fore-

casted boundary conditions from the perfect initial field and the previous dynamical field. The third is to combine the advected boundary condition from the previous dynamical field and the statistically forecasted boundary condition from the perfect initial field. They all show significant improvement over the persistent dynamical forecast (25 percent vs. 60 percent error after two weeks). The price we pay for the improvement is the additional statistical parameters required for the combination. Whether there is any better scheme to estimate the boundary condition for the particular error model chosen is suggested for future investigations.

However, the success of any model depends on its ability to work with the real data. Our models, though they work satisfactorily with the simulation data, have to take up this challenge too. The statistics needed for the optimal combination are presently computed from the simulation data in which a verification data is available. The averages of these statistics from many realizations will be the starting point for working with the real data. The final statistics to be used are the ones that work best with the largest practical real data set. They may have to be computed from the real ocean data if necessary.

The Harvard Open Ocean Modelling Group, who works jointly with the MODE and POLYMODE Group, has the priviledges of accessing the real ocean data. These real ocean data will be used to verify the results obtained and evaluate the models con-

structed. The models have to go through many iterations before they can physically realistically simulate the real ocean. The models can then be used to study some of the critical dynamical questions in the oceans--"What are the sources of the eddies that populate the North Atlantic gyre, how are they coupled to mean flow of the general circulation, and what is their contribution to the physics that controls the general circulation?" (Robinson (24)) They are useful in assessing the dynamical hypotheses, interpreting the field data, and planning future open ocean dynamics experiments: where and when to collect data in an open ocean in the most efficient way so that the forecast error is within certain tollerance limits. This is a complicated and challenging problem. There is still a lot of work waiting to be done.

BIBLIOGRAPHY

(1) Arakawa, A., 1966. Computational design for long-term numerical integration of the equations of fluid motion: Two-dimensional incompressible flow. Part I. J. Comp. Phys., 1, 119-143.

(2) Bengtsson, L. and N. Gustavsson, 1971. An experiment in the assimilation of data in dynamical analysis. Tellus, 23, 328-336.

(3) Bengtsson, L. and N. Gustavsson, 1972. Assimilation of non-synoptic observations. Tellus, 24, 383-399.

(4) Bengtsson, L., 1975. 4-dimensional assimilation of meteorological observations. GARP Publication Series #15.

(5) Bennett, A. and P. Kloeden, 1978. Boundary conditions for limited area forecasts. J. Atm. Sci., 35, 990-996.

(6) Box, G.E.P. and G.M. Jenkins, 1970. Time series analysis: Forecasting and control. San Francisco, Holden-Day.

(7) Charney, J.G., 1948. On the scale of atmospheric motions. Geophys. Publ., 17(2), 3-17.

(8) Courant, R., K. Friedrichs, and H. Lewy, 1967. On the partial difference equations of mathematical physics. IBM Journal, 215-234.

(9) Fix, G.J., 1975. Finite element models for ocean circulation problems. SIAM J. App. Math., 29, 371-387.

(10) Gandin, L.S., 1965. Objective analysis of meteorological fields. (Translated from Russian.) Israel Program for Scientific Translations, Jerusalem.

(11) Ghil, M., M. Halem, and R. Atlas, 1979. Time-continuous assimilation of remote-sounding data and its effect on weather forecasting. Monthly Weather Rev., 107(2), 140-171.

(12) Haidvogel, D.B., A.R. Robinson and E.E. Schulman, 1980. The accuracy, efficiency, and stability of three numerical models with application to open ocean problems. J. Comp. Phys., 34(1), 1-53.

(13) Hirsh, J.E., 1973. Advanced methods in numerical oceanography. Ph.D. thesis, Harvard University.

(14) Jespersen, D.C., 1974. Arakawa's method is a finite-element method. J. Comp. Phys., 16, 383-390.

(15) Knuth, D.E., 1973. The art of computer programming, Volume 3 -- Sorting and Searching. Addison-Wesley, Reading, MA., 84-95.

(16) McWilliams, J.C. and G. Flierl, 1976. Optimal, quasi-geostrophic wave analyses of MODE array data. Deep-Sea Res., 23, 285-300.

(17) Pereyra, V., 1967. Accelerating the convergence of discretization algorithms. SIAM J. Num. Anal., 4, 508-533.

(18) Petersen, D.P. and Middleton, D., 1964. Reconstruction of multidimensional stochastic fields from discrete measurements of amplitude and gradient. Info. and Cont., 7, 445-476.

(19) Petersen, D.P. and D. Middleton, 1965. Linear interpolation, extrapolation, and prediction of random space-time fields with a limited domain of measurement. IEEE Trans. Info. Th., IT-11, 18-30.

(20) Petersen, D.P., 1965. Recursive formulas for sequential estimation. IEEE Trans. Info. Th., IT-11, 589-590.

(21) Shapiro, R., 1970. Smoothing, filtering, and boundary effects. Rev. Geophys. Space Phys., 8, 359-387.

(22) Shapiro, R., 1971. The use of linear filtering as a parameterization of atmospheric diffusion. J. Atm. Sci., 28, 523-531.

(23) Sundström, A., 1969. Stability theorems for the barotropic vorticity equation. Mon. Wea. Rev., 97(4), 340-345.

(24) Robinson, A., 1976. Eddies and ocean circulation. Oceanus, 2-17.

(25) Robinson, A., 1976. Numerical modeling and global ocean forecasting. Science, Technology, and the Modern Navy, 481-496.

(26) Robinson, A.R. and D.B. Haidvogel, 1980. Dynamical forecast experiments with a barotropic open ocean model. To appear in J. Phys. Oceanogr.

(27) Robinson, A., and W. Simmons, 1980. A new dimension in physical oceanography. Oceanus, 40-51.

(28) Rutherford, I.D., 1972. Data assimilation by statistical interpolation of forecast error fields, J. Atm. Sci., 29, 809-815.